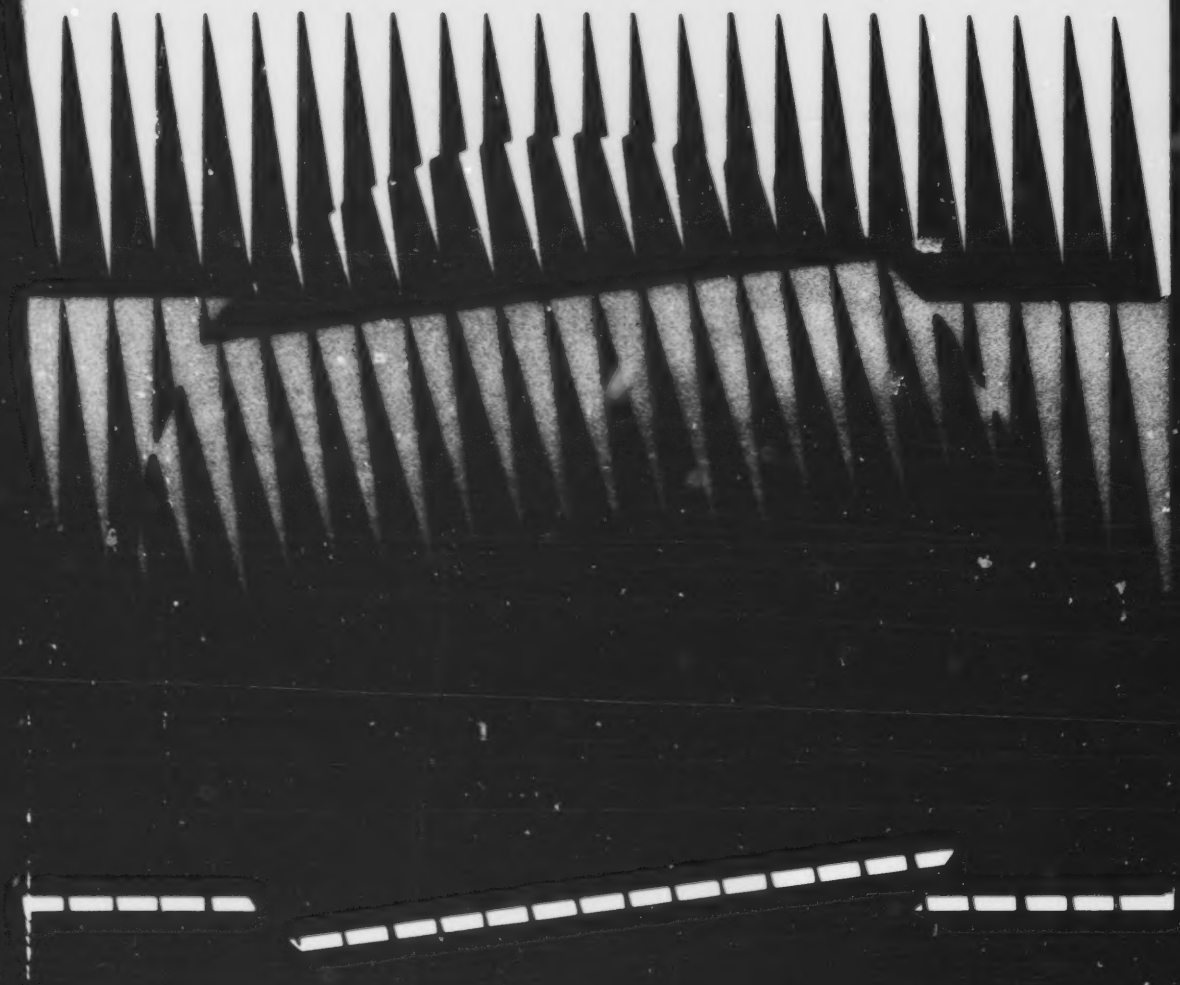
Canadian Bulletin of Fisheries and Aquatic Sciences 213

ECOSYSTEM THEORY FOR BIOLOGICAL OCEANOGRAPHY

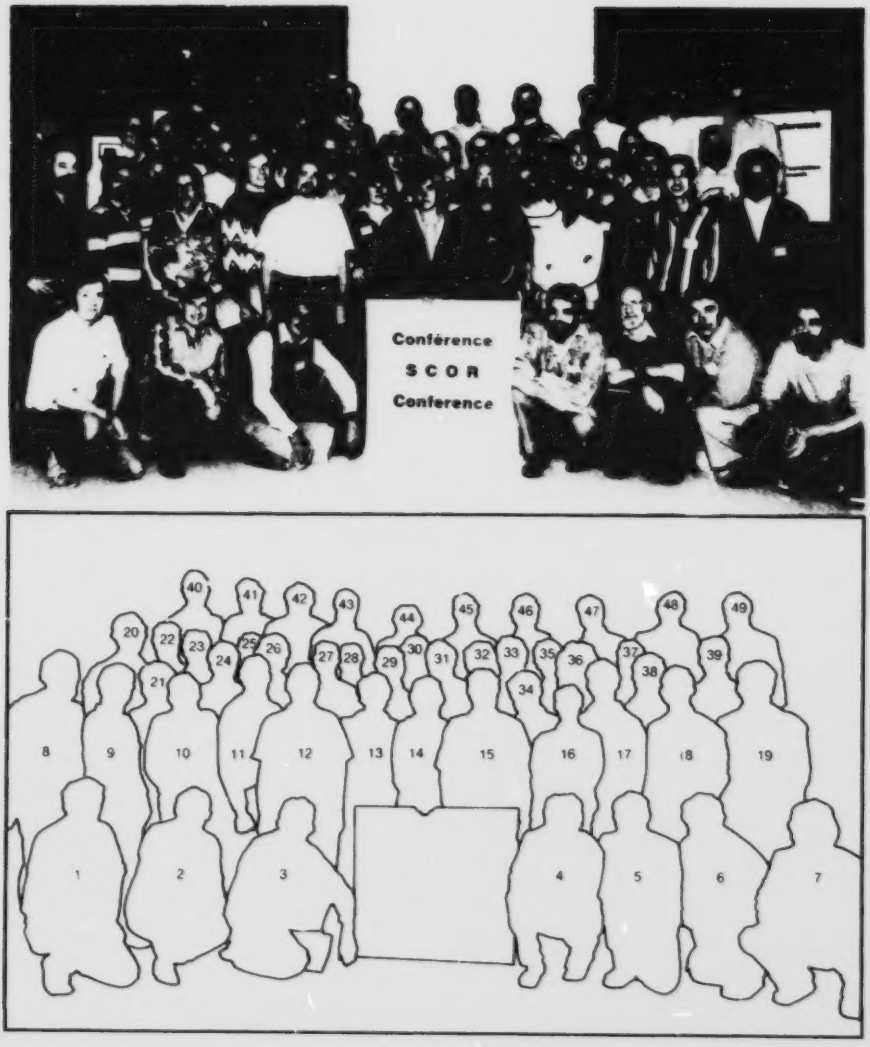
Edited by Robert E. Ulanowicz and Trevor Platt



Fisheries and Oceans Pêches et Océans

Canada

ECOSYSTEM THEORY FOR BIOLOGICAL OCEANOGRAPHY



1 — Richard Haedrich, 2 — Peter Allen, 3 — Victor Smetacek, 4 — Jordi Flos, 5 — Lynn Trainor, 6 — Lionel Johnson, 7 — Marcel Fréchette, 8 — Michael Conrad, 9 — Rene Ulloa, 10 — Gilberto Rodriguez, 11 — Robert Ulanowicz, 12 — Louis Legendre, 13 — Clarice Yentsch, 14 — Marta Estrada, 15 — Serge Frontier, 16 — Shubha Sathyendranath, 17 — Serge Demers, 18 — Hironori Hirata, 19 — Patricio Bernal, 20 — Curtis Davis, 21 — Juan Tundisi, 22 — Michael Kemp, 23 — Yves De La Fontaine, 24 — Henri Adan, 25 — James Kay, 26 — Pierre Lasserre, 27 — Ramon Margalef, 28 — Trevor Platt, 29 — Jacqueline McGlade, 30 — Donald Mikulecky, 31 — Robert Rosen, 32 — Peter Wangersky, 33 — Charles Yentsch, 34 — Kenneth Denman, 35 — David F. Smith, 36 — Isabelle Lamontagne, 37 — William Calder, 38 — Walter Boynton, 39 — John Field, 40 — Steven Cousins, 41 — William Silvert, 42 — Claude Joiris, 43 — William Leggett, 44 — Michael Fasham, 45 — Bernard Patten, 46 — Marlon Lewis, 47 — Vincent Gallucci, 48 — Fredrik Wulff, 49 — Alexander Zotin. *Not pictured*: Jean-Yves Bellavance, Bruce Hannon, Patricia Lane, Jeffery Watson.

CANADIAN BULLETIN OF FISHERIES AND AQUATIC SCIENCES 213

Ecosystem Theory for Biological Oceanography

(Proceedings of a Symposium sponsored by SCOR, NSERC, NSF, UNESCO, and the U.S. Office of Naval Research held at Laval University, Ste. Foy, Quebec, March 16–23, 1984.)

EDITED BY

ROBERT E. ULANOWICZ

University of Maryland
Chesapeake Biological Laboratory
Solomons, Maryland 20688, USA

AND TREVOR PLATT

Department of Fisheries and Oceans
Marine Ecology Laboratory
Bedford Institute of Oceanography
Dartmouth, Nova Scotia B2Y 4A2

DEPARTMENT OF FISHERIES AND OCEANS
Ottawa 1985

The *Canadian Bulletins of Fisheries and Aquatic Sciences* are designed to interpret current knowledge in scientific fields pertinent to Canadian fisheries and aquatic environments.

The *Canadian Journal of Fisheries and Aquatic Sciences* is published in annual volumes of monthly issues. *Canadian Special Publications of Fisheries and Aquatic Sciences* are issued periodically. These series are available from authorized bookstore agents and other bookstores, or you may send your prepaid order to the Canadian Government Publishing Centre, Supply and Services Canada, Ottawa, Ont. K1A 0S9. Make cheques or money orders payable in Canadian funds to the Receiver General for Canada.

Scientific Information and Publications Branch

J. Watson, Ph.D.	Director and Editor-in-Chief
Johanna M. Reinhart, M.Sc.	Scientific Evaluation and Information Systems
J. Camp / G. J. Neville / Susan D. Scale	Editorial and Publishing Services

Editorial Office: Department of Fisheries and Oceans
Scientific Information and Publications Branch
200 Kent Street, 12th Floor
Ottawa, Ontario, Canada K1A 0E6

Typesetter: K. G. Campbell Corporation, Ottawa, Ontario
Printer: Kromar Printing Ltd., Winnipeg, Manitoba
Cover Design: Ludvic Saleh & Associates, Ottawa, Ontario

©Minister of Supply and Services Canada 1985

Available from authorized bookstore agents, other bookstores
or you may send your prepaid order to the
Canadian Government Publishing Centre
Supply and Services Canada, Ottawa, Ont. K1A 0S9.

Make cheques or money orders payable in Canadian funds
to the Receiver General for Canada.

A deposit copy of this publication is also available
for reference in public libraries across Canada.

Canada: $14.95	Cat. No. Fs 94-213E
Other Countries: $17.95	ISBN 0-660-11900-5
	ISSN 0706-6503

Price subject to change without notice

Correct citation for this publication:

ULANOWICZ, R. E., AND T. PLATT [ED.]. 1985. Ecosystem theory for biological oceanography. Can. Bull. Fish. Aquat. Sci. 213: 260 p.

Contents

Abstract / Résumé ..... vii–viii
Background and Acknowledgements ..... ix–x
Guide to the Contents. Robert E. Ulanowicz and Trevor Platt ..... xi–xiii

I. Thermodynamics

Ecology, Thermodynamics, and Self-Organization: Towards a New Understanding of Complexity. P. M. Allen 3–26
Thermodynamics and Growth of Organisms in Ecosystems. A. I. Zotin 27–37

II. Statistical Mechanics

Statistical Mechanics in Biology — Applications to Ecology. L. E. H. Trainor 41–51

III. Ataxonomic Aggregations

Structure of the Marine Ecosystem: Its Allometric Basis. Trevor Platt .. 55–64
Size and Metabolism in Natural Systems. William A. Calder III 65–75
The Trophic Continuum in Marine Ecosystems: Structure and Equations for a Predictive Model. S. H. Cousins 76–93

IV. Flow Analysis

Ecosystem Flow Analysis. Bruce Hannon 97–118
Energy Cycling, Length of Food Chains, and Direct Versus Indirect Effects in Ecosystems. Bernard C. Patten 119–138
Flow Analysis of Materials in the Marine Euphotic Zone. M. J. R. Fasham 139–162
Network Thermodynamics in Biology and Ecology: An Introduction. Donald C. Mikulecky 163–175

V. Information Theory

The Statistical Basis of Ecological Potentiality. Michael Conrad 179–186
Information Theory and Self-Organization in Ecosystems. H. Atlan ... 187–199
From Hydrodynamic Processes to Structure (Information) and from Information to Process. Ramon Margalef 200–220
Information and Complexity. Robert Rosen 221–233

VI. Working Group Reports

I. Hypothesis Testing and Sampling Design in Exploited Ecosystems. W. C. Leggett (Chairman) 237–240

II. Ecosystem Theory in Relation to Unexploited Marine Ecosystems. J. G. Field (Chairman) 241–247

III. Possible Holistic Approaches to the Study of Biological–Physical Interactions in the Oceans. L. Legendre (Chairman) 248–253

IV. Technoiogical Developments to Implement Theory into Biological Oceanography. K. Denman (Chairman) 254–258

V. The Design of Large-Scale Cooperative Experiments. P. A. Bernal (Chairman) 259–260

Abstract

ULANOWICZ, R. E., AND T. PLATT [ED.]. 1985. Ecosystem theory for biological oceanography. Can. Bull. Fish Aquat. Sci. 213: 260 p.

Holistic descriptions of marine ecosystems offer an alternative to characterizing biotic communities in terms of coupled process models. Most quantitative narratives of whole ecosystem organization and development draw from five subject areas: *Thermodynamics* provides the phenomenological and macroscopic perspective necessary to perceive whole system behavior. *Statistical mechanics* is an historical example of how microscopic observations and hypotheses may be reconciled with their macroscopic counterparts. *Ataxonomic aggregations* of organisms, such as classifications according to particle size or metabolic rate, rely less upon "microscopic" (i.e., taxonomic) features of the ecosystem and could be more appropriate elements with which to build holistic theories in marine ecology. *Flow analysis* is quantitative theory germane to the study of how the parts of an ecosystem directly and indirectly affect each other within the context of the entire system. Lastly, *information theory* is a formalism capable of bridging and ultimately unifying the preceding four disciplines.

Holistic considerations of ecosystem behavior tend, at this early stage, to be highly abstract. The existing theories, nonetheless, have practical implications for existing biological programs. Various holistic hypotheses may be tested in both exploited and non-exploited marine ecosystems. Macroscopic concepts can promote a better understanding of biological–physical interactions and suggest the development of new technological instrumentation and methods. Finally, whole community descriptions beg for the design of new, large-scale, cooperative experiments in biological oceanography.

Résumé

ULANOWICZ, R. E., AND T. PLATT [ED.]. 1985. Ecosystem theory for biological oceanography. Can. Bull. Fish Aquat. Sci. 213: 260 p.

Les descriptions holistiques des écosystèmes marins offrent une autre possibilité pour la caractérisation des communautés biotiques sous forme de modèles de processus jumelés. La plupart des descriptions quantitatives de l'organisation et du développement d'un écosystème complet relèvent de cinq domaines: la *thermodynamique* fournit la perspective macroscopique et phénoménologique nécessaire à la perception du comportement du système entier; la *mécanique statistique* constitue un exemple historique de la manière dont les observations microscopiques et les hypothèses peuvent être conciliées avec leurs analogues macroscopiques; les *groupements non taxonomiques* des organismes, comme la classification selon la taille des particules ou le taux métabolique, s'appuient moins sur les caractéristiques «microscopiques» (c.-à-d. taxonomiques) de l'écosystème et pourraient constituer des éléments plus appropriés pour l'élaboration de théories holistiques en écologie marine; *l'analyse du déroulement* est une théorie quantitative convenant à l'étude de l'interaction directe et indirecte ente les parties d'un écosystème dans le contexte de l'écosystème entier; finalement, la *théorie de l'information* est un formalisme capable de relier et, en fin de compte, d'unifier les quatre disciplines précédentes.

Les considérations holistiques relatives au comportement de l'écosystème ont tendance, à ce stade, à être très abstraites. Néanmoins, les théories existantes ont des portées pratiques pour les programmes biologiques actuels. On peut vérifier diverses hypothèses holistiques dans des écosystèmes marins exploités et non exploités. Par ailleurs, les concepts macroscopiques peuvent encourager une meilleure compréhension des interactions biologiques–physiques et inspirer la mise au point de nouveaux instruments et méthodes. Enfin, pour la description de communautés entières il faut concevoir de nouvelles expériences coopératives à grande échelle en océanographie biologique.

Background and Acknowledgements

This bulletin is the latest of three related volumes to issue from a group that traces its origins to the Scientific Committee on Oceanic Research (SCOR), Working Group Number 59, "Mathematical Models in Biological Oceanography". SCOR charged this body with suggesting mathematical methods in marine ecology useful for the design of research programs and for the treatment of biological oceanographic data.

The report of SCOR WG59 appeared as a book (Platt et al. 1981) which bore the same title as the Working Group. The committee acknowledged the utility and encouraged the use of mechanistic models to quantify very simple biological systems, such as those consisting of a single process or species. However, the group advised that the application of differential equations to simulate coupled processes or entire biotic communities should be undertaken with strict attention on the part of the modeler to the significant limitation inherent in such an approach.

Conscientious criticism of any endeavor obliges those making the judgement to outline what they perceive to be constructive and fruitful alternative paths of inquiry. WG59 cited two primary directions in which they believed new investigations should proceed. First, they claimed that to understand biological oceanographic systems, it is necessary to have at least as much information on the fluxes as on the biomasses. To emphasize the need for measurements of "flows of energy and materials in marine ecosystems" SCOR WG59 sponsored, on the same topic, a NATO Advanced Research Institute whose proceedings have recently appeared (Fasham 1984).

The second major recommendation of the committee was to promote the development of holistic approaches to the study of marine ecosystems. The committee identified five disciplines which showed *potential* for helping to define and measure community-level properties in biological oceanographic systems: Thermodynamics, Statistical Mechanics, Input−Output Analysis, Information Theory, and Ataxonomic Aggregations. They proposed to SCOR that a conference be held to bring leading theoreticians in these five fields into contact with a representative group of biological oceanographers who would be favourably disposed to incorporate such theoretical notions into their work if a convincing case for their utility could be made.

In parallel with this new direction, WG59 was dissolved, and WG73, "Ecosystems Theory in Relation to Biological Oceanography", was established and provided with seed money for the proposed symposium. Working Group Chairman, Kenneth Mann appointed an Organizing Committee consisting of himself, Co-Chairman Trevor Platt, and Louis Legendre with Robert Ulanowicz as head, to attract co-sponsors and to plan the meeting. The Natural Sciences and Engineering Research Council of Canada, the Biological Oceanography Program of the U.S. National Science Foundation, the U.S. Office of Naval Research, and the United Nations Educational and Scientific Council were generous in their support as co-sponsors of the meeting.

The colloquium took place at the Forêt Montmorency Conference Center, Ste. Foy, Quebec, on March 16−23, 1984. The Center is run by Laval University, which contributed significant non-monetary support to the project. Local arrangements were superbly directed by Prof. Louis Legendre, who in turn was ably assisted in this task by Isabelle Lamontagne, Jean-Yves Bellavance, Yves Delafontaine, and Marcel Frechette of the Groupe interuniversitaire de recherches océanographiques du Québec (GIROQ).

The editors also wish to thank Mrs. Elizabeth Tidmarsh, Executive Secretary of SCOR, for the vital part she played in handling the administrative affairs for the meeting.

References

FASHAM, M. J. R. [ED.] 1984. Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY. 733 p.

PLATT, T., K. H. MANN, AND R. E. ULANOWICZ. 1981. Mathematical models in biological oceanography. Monographs on oceanographic methodology 7, UNESCO Press, Paris. 157 p.

Guide to the Contents

This work is intended as an appeal to Biological Oceanographers to entertain a more global and holistic perspective on the marine ecosystems they study. How to adopt a macroscopic outlook on ecology cannot be elaborated here in finished detail, but there is certainly precedent for large-scale vision in science. The description of macroscopic phenomena is at the core of *thermodynamics*, the discipline upon which both Einstein and Eddington have bestowed primacy over all the other sciences.

While most treatments of thermodynamics in biology center around the implications which the two fundamental laws have for living systems, a different approach is adopted here. The emphasis in this book is upon the phenomenological method and the benefits it can bring to the early stages of any investigation. The lead article by Allen is a summary of some of the recent intriguing insights that have evolved out of thermodynamics. Allen urges biological oceanographers to regard marine ecosystems as dissipative structures and proceeds to portray some of the colorful and exciting attributes of dissipative systems that are missing from mechanical analogs. Zotin follows with a classical example of phenomenological induction. He begins with a large assortment of empirical equations for growth and gathers them into a unified description of this fundamental and ubiquitous process.

The development of thermodynamics proceeded for longer than half a century before any serious attempt was made to reconcile the body of macroscopic principles with events at the microscopic, or atomic scale. This effort, only partially successful, embodied the discipline of *statistical mechanics*, and it is only reasonable to ask whether the methods of statistical mechanics are appropriate to bridging the gap between ecological measurements made at the scale of the organism or population (i.e., the "microscopic" domain) and those observations made on the macro level of the whole ecosystem. The consensus was less than encouraging, as witnessed by the fact that two of the three invited sepakers on this topic eventually gravitated towards issues that were better placed in other sections of the conference. Trainor, however, addresses the question head-on and is refreshingly frank in his assessment of the prospects for applying statistical mechanics to ecology.

It becomes clear, therefore, that little is to be gained by retreading old pathways, and that innovation is needed if an adequate phenomenological description of marine ecosystems is to be accomplished.

One of the little-heralded but important attributes of classical thermodynamics is that it forms a self-consistent body of description without any necessary recourse to the details of microscopic events. One may teach an entire course in thermodynamics without ever once mentioning the words "atom" or "molecule". It is possible to overemphasize this apparent independence of macroscopic narrative from microscopic detail, but the autonomy of large-scale description should cause one to reassess the terminology in which traditional ecology is cast. For example, most conceptual models of ecosystems start from a basis of taxonomy. But taxonomic differentiation is based on "microscopic" features or organisms, or sometimes even on genetic or molecular distinctions. How appropriate, then, are taxonomic classifications to the desired macroscopic treatment of ecosystems? Might not some *ataxonomic aggregations* be more effective as elements with which to describe whole-community behavior? Perhaps organism size or metabolic rates are more natural characteristics upon which to define ecosystem components. Such descriptors possess actual physical dimensions, and this attribute imposes a degree of coupling between any two characteristics with overlapping dimensionality, i.e., an allometric relationship.

Biological Oceanography has pioneered the description of marine ecosystems in terms

of ataxonomic aggregations and allometry. The papers by Platt and by Cousins review this work. That of Calder brings the further perspective of similar developments in terrestrial ecology.

The power of thermodynamics lies in its universality. But to work with generalities and still say something meaningful, it becomes necessary to focus upon a category of real phenomena that is all-encompassing, yet readily measurable. Enter the recommendation by WG59 to emphasize flow processes. Flows may be quantified in any non-equilibrium process, and the study of flows in ecosystems has commanded progressively more attention over the past decade.

It was Hannon who first introduced Input–Output Analysis, the vector analysis of economic flows, into ecology; and here he employs *flow analysis* to compute the total "cost" of producing any ecological product in terms of all its antecedent processes. These intensities are calculated for three marine food webs, and the significance of those measures is discussed. Patten has led the way in expanding the use of flow analysis in ecosystems studies. In this symposium he demonstrates the importance of energy cycling and the significance of indirect effects in most ecological networks. Fasham follows with a thought-provoking application of flow-analysis to the ecosystem of the euphotic zone of warmwater marine systems. He demonstrates that the recycle of DOC through the microbes is probably a very significant contribution to the production of higher trophic-level species. In the last paper on flow networks Mikulecky turns the reader's attention back towards the first section with his review of network thermodynamics, the attempt to marry circuit analysis with near-equilibrium thermodynamics.

If thermodynamics prescribes how one is to observe macroscopic systems, and flows represent appropriately general objects to measure; then one may ask what formalism might be employed to tie these aspects together? A clue is found in the history of thermodynamics (or statistical mechanics to be more precise), where Boltzmann developed the precursor of modern *information theory*. Indeed, Conrad finds the information–theoretic concept of conditional entropy to be an excellent tool with which to expound upon the adaptability that allows ecosystems and other living entities to maintain their existences. Atlan argues that information theory itself is insufficient to formalize what happens in living systems; that it is also necessary to define and measure "meaning" in the context of a system. He provides a simple but exciting numerical example of how meaning might arise in a stochastically self-organizing system. Among other things, Margalef is known for having introduced information theory into the study of marine ecosystems. In this forum he uses his experience with plankton patches to attempt to unify all the sections of the conference: energetics with information theory in a non-deterministic scenario; ataxonomic spectral representations with the connectivity of flows via the theory of fractals. Rosen fittingly ends the series of contributions with an epistemological analysis of information that ultimately leads him independently to re-iterate Allen's call for a non-Newtonian theory of ecosystems. Ecosystems and organisms are not machines; and, as SCOR WG59 implied, cannot be modeled in mechanistic fashion in any but the most restricted circumstances.

Most of the contributors are not experts in marine ecology, as perusal of these articles will reveal. Therefore, it fell to the participants in the conference to interpret the ideas and approaches presented at the meeting into concrete recommendations for action by biological oceanographers. This translation was attempted by five working groups, which met each afternoon of the symposium to draft reports to SCOR on the subjects of how to test these hypotheses in (1) exploited and (2) non-exploited ecosystems, (3) how do the concepts help in understanding biological–physical interactions, (4) what technological developments are necessary to implement the ideas, and (5) what large-scale cooperative

experiments might be undertaken in the light of the theories presented. These reports round out the proceedings.

The decades of the 50's and 60's saw brilliant advances in genetics and molecular biology. It is the opinion of the organizers of this conference that the time is right for equally exciting discoveries in macrobiology. The editors hope that these proceedings will be a significant impetus in moving biological oceanographers along this road.

ROBERT E. ULANOWICZ
Solomons, Maryland

TREVOR PLATT
Dartmouth, Nova Scotia

I. THERMODYNAMICS

Ecology, Thermodynamics, and Self-Organization: Towards a New Understanding of Complexity

P. M. ALLEN

CP. 231, Chimie Physique II, Université Libre de Bruxelles, Brussels 1050, Belgium

Introduction

The science of ecology is remarkable in that it aims specifically for an understanding of a whole system. It accepts as given the fact that the "whole is more than the sum of its parts", and sets out to reveal and comprehend the parts and the whole.

This, however, is not an easy task in that ecosystems are inevitably composed of a whole series of interdependent populations. The characteristics of these populations result from their mutual oevolution, together with their interactions with their environment. The science of such systems — of complex systems — is only beginning to be understood (see also Rosen, this volume).

The first phase of any science is that of "classification". Thus, the diversity of nature must be catalogued, and any recurrent patterns or regularities noted. This endeavour has led to different representations of ecosystems ranging from the detailed compartmentalization of taxonomic classes present in the system to much more aggregate depictions as "trophic levels", or as classes of organism size or type. The "ecosystem" is then represented by the flows of different materials (carbon, nitrogen, etc.) or of energy between these groupings.

All this is one phase of scientific endeavour, and it tells us what we must try to "explain" on the basis of some more profound, or lower level invariances, which would permit us to "predict" or "expect" that certain patterns would be observed in specific environments.

For example, if we examine the energy flows and stocks in a particular ecosystem, we find a diagram such as that shown in Fig. 1. Such a scheme is already very interesting and represents succinctly a considerable amount of information concerning the structure and functioning of the system. However, the questions which any eco-science should try to address are deeper than this. For example, why consider only these boxes? Why are there not others? Why are there arrows between some boxes but not between all? Why do the flows and stocks have even approximately these values? What is driving the system — or in other words, where is the *vitality* of the system? Briefly, we may ask what is it that makes the system what it is; and in the case that we could identify this factor, then could we say what we would expect of a system in other external circumstances? Also, and very importantly, if we decided to intervene in the ecosystem, either to harvest it, or to pollute it (or as generally happens, to do both), then what is likely to happen?

These are key questions that should be addressed. But having said that, the problem is first to decide on what basis an answer could be given, since neither the pristine harmony of classical mechanics, nor the disordered states of equilibrium thermodynamics offer a suitable paradigm for the budding, interdependent organisation of living systems. Recently, however, new concepts have emerged from the study of open systems that are maintained far from thermodynamic equilibrium. These ideas offer us a new basis from which to understand the origin and evolution of structure and organisation, and can be used to describe a process which corresponds much more closely to the creative development that we know characterises living systems. (Nicolis and Prigogine 1977).

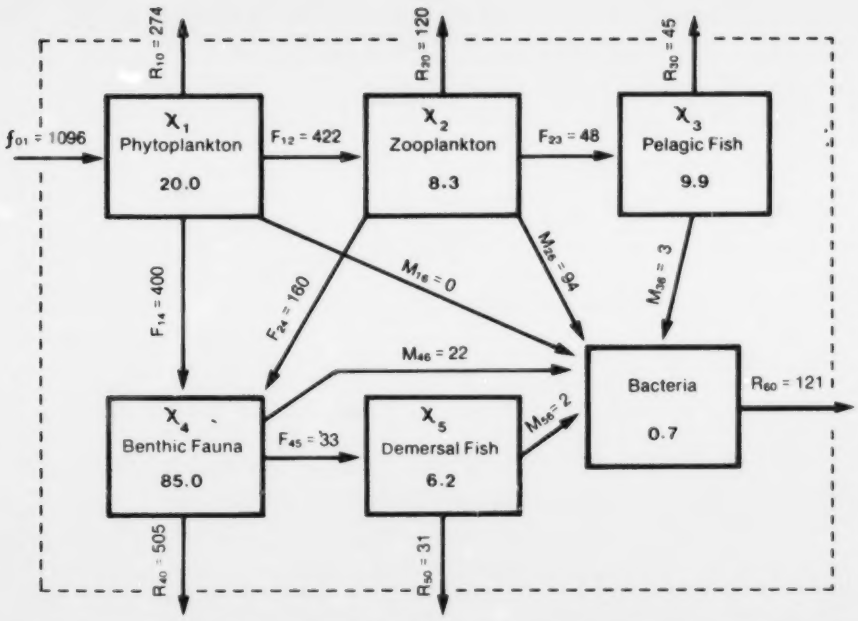


FIG. 1. Block diagram of energy flow for the English Channel. Standing crops are in kilocalories per square metre and energy flows are in kilocalories per square metre per year (from Brylinsky 1972).

These new ideas pose a new type of "explanation" of phenomena. The scientific view of "explanation" has long been associated with understanding macro-phenomena in terms of some micro-description — a typically reductionist point of view. The alternative was always thought to be simple "phenomenology" or description, where certain relations between macro-variables were found to hold and, therefore, constituted a consistent body of mutual "explanations". This latter path, although useful, seems to me to be much less profound than the former. However, as we shall see, open systems that are far from thermodynamic equilibrium can give rise to macro-structure which has a complex and non-unique relationship to the micro-description of its elements. In this way we find that such systems can truly give rise to entities which are more than the sum of their parts. What may begin as "reductionism" ends up in the "elevation" of the material components of matter, and to a realization that the "explanation" of a given macroscopic situation results from the mutual consistency of two levels of description — the micro and the macro levels.

Dissipative Structures – Models of Complexity

In order to illustrate as clearly as possible the surprising new aspects of the ideas to be presented, let us turn back briefly to classical physics, and to the manner in which the passage of time was considered to "mark" a system. For example, in Fig. 2 we see how a system of gas molecules evolves when, after being enclosed in half the box by a partition, they are allowed to fill the whole space.

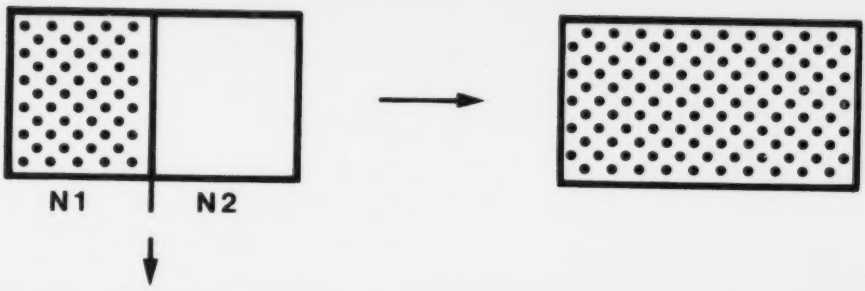


FIG. 2. An irreversible movement to uniformity follows the removal of a partition dividing the space in two.

The passage of time is marked, after removal of the partition, by an irreversible movement towards uniformity. After some interval, if we take random readings of the number of molecules on the left and right hand sides of the box, we shall always find approximately the same number. Even if we had started with the opposite initial condition, after some time we would observe precisely the same result. The precise initial condition is of no importance; the evolution is always towards uniformity, that is, maximum disorder. The reason is very simple. The number of possible molecular combinations, P, which can give rise to $N1$ on one side and $N2$ on the other is:

$$(1) \quad P = \frac{N!}{N1!N2!}, \qquad N = N1 + N2,$$

and P is maximum when $N1 = N2$. Thus, if we suppose that all microstates are of equal probability, then we shall simply observe a "movement" from an improbable initial macrostate to the most probable, i.e., uniformity.

Boltzmann (1872) was able to go further and to identify these molecular combinations with the "Entropy" through the relation:

$$(2) \quad S = k \log P,$$

and hence was able to link the macrolevel of entropy and the second law with the microlevel of molecular motion and to show that the second law implies ever growing molecular disorder in the universe, increasing banality and the destruction of pre-existing order.

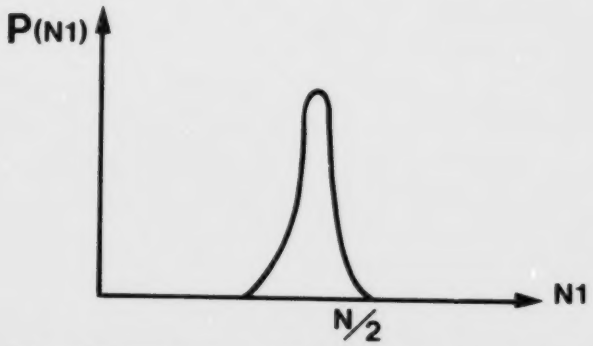


FIG. 3. The number of molecular combinations corresponding to a particular value of $N1$.

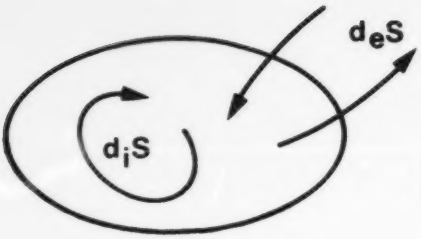


FIG. 4. The entropy change of an open system has both internal and external contributions.

However, I must underline the fact that this description of the relationship between levels pertains only to an isolated system made up of weakly coupled particles. If we turn to open systems, and allow the possibility of strong interactions between elements, then the whole picture can change.

In fact, the second law says only that in an open system with fluxes of matter and energy across its boundaries the contribution to entropy production coming from the processes occurring within the system will always be positive or zero. However, the entropy flow to the system through the walls, d_eS, if it were sufficiently negative, could more than offset the internal dissipation, and lead to a negative net change in entropy for the system.

For a system maintained just out of the thermodynamic equilibrium by, for example, a weak temperature gradient, the steady state attained by the system is unique, and as near as possible to the disordered, or equilibrium state. This was shown by Prigogine (1947) in his theorem concerning minimum entropy production.

In a system with only "weak" interactions between its elements and subjected to stronger disequilibria, still nothing remarkable may occur. However, in a system with some degree of strong coupling between its elements, when its disequilibrium reaches a certain intensity, then many amazing and surprising things can happen.

In order to illustrate this let us briefly describe the behaviour of a particular chemical system which has served as a "prototype" for understanding these new ideas. It is known as the "Brusselator", because of the intensive study it has received by the group at Brussels.

It consists of a simple, fixed, nonlinear reaction mechanism,

$$\begin{aligned} &A \rightarrow X \\ &B + X \rightarrow Y + D \\ &2X + Y \rightarrow 3X \\ &X \rightarrow E, \end{aligned} \tag{3}$$

where A and B feed the reaction, D and E are produced by it, and X and Y are intermediates. Let us suppose further that X is red in colour, and Y is blue.

The kinetic equations for this reaction scheme are very simple to write, and we assume that the products of reaction E and D are removed to avoid the occurrence of a back reaction:

$$\begin{aligned} \frac{dX}{dt} &= A - BX + X^2Y - X \\ \frac{dY}{dt} &= BX - X^2Y. \end{aligned} \tag{4}$$

There exists a stationary state solution to these equations: $X' = A$; $Y' = B/A$. However, this system can be driven far from thermodynamic equilibrium by increasing A and B; and if this is done, many different possibilities arise. For example, if we stir the reaction, then at a certain critical reaction rate, instead of the system being uniform (a homogeneous mixture of red and blue, of X and Y) it suddenly begins to oscillate steadily from all red to all blue and back, in a perfectly rythmic manner. Even if perturbed momentarily, it will return to this particular, stable beat. The random, incoherent, blind movements and reactions of the molecules is abruptly transformed into disciplined, coherent, coordinated behaviour worthy of a good choir!

But this outcome is only one of many possibilities. If we take a system which is not constantly homogenized by stirring, then as we move further from equilibrium, all sorts of spatial and spatio-temporal structures can appear spontaneously: from simple left/right inhomogeneities, to expanding spiral waves of various well defined dimensions, to moving or stationary bands of red and blue — a whole bundle of different possibilities. (See Fig. 5).

This process of self-organization is a remarkable phenomenon which strikes at the heart of some of our deepest preconceptions concerning physical systems. For example, if we take a particular spatial structure, then at the interface of "red" and "blue" there will clearly be fluxes of X and Y caused by the concentration gradients. Our normal reaction would be to say that they are "explained" by the "forces" that must exist between the zones. But in fact these forces themselves are generated by the spatial structure of which the interface is a part, and which in turn reflects the fluxes that are occurring in the system. If, for example, the coefficient of diffusion were modified, or the temperature, then the spatial structure itself would change or perhaps even disappear. In this sense, the "cause" of this particular structure is the precise values of the fluxes, which in their turn, according to our simple preconceptions, result from this structure. Clearly, the circularity of the apparent "causation" is showing up some weakness in our way of thinking about things.

In reality, a "dissipative structure" is an entity which has as mutually dependent facets the flows and spatial structures that characterize it. Interference with one will modify both through a cascade of feedback processes. We see that our fundamental questions concerning the ecosystem model characterized in Fig. 1 are answered in a surprising way by the Brusselator, because there we behold a system that has created its own "boxes" and "arrows". Furthermore, we see that the "Modeller's Nightmare", i.e., the fact that complex systems evolve structurally (new boxes, new arrows) is quite clearly part of the behaviour of a dissipative structure. A particular type of behaviour, homogeneous temporal oscillation, moving parallel bands, etc. can spontaneously change to a qualitatively different one. If we had been rash enough to model the system on the basis of its particular macrobehaviour at the earlier time, then suddenly our model would fail to describe what was occurring.

Also, we come upon the dilemma that faces any ecologist trying to understand the system before him. We can "track" the energy flow in the Brusselator, or make balance equations (accountancy) for particular materials (carbon, nitrogen etc.); but these always only *indicate* or *reflect* the structure that had appeared in the system, and do not explain it, nor predict when some new structure may emerge.

The "explanation" behind a particular "structure/flow" pattern lies in the history of instabilities it has traversed, and especially in the stability or instability of the structure at the moment we are observing it. All that is necessary for such an entity to persist is that it be stable with respect to the perturbations to which it is being subjected. If we plot the value of system property X at a point r of our system, $X(r)$, then what we find is an "evolutionary tree" of possible behaviours for our system.

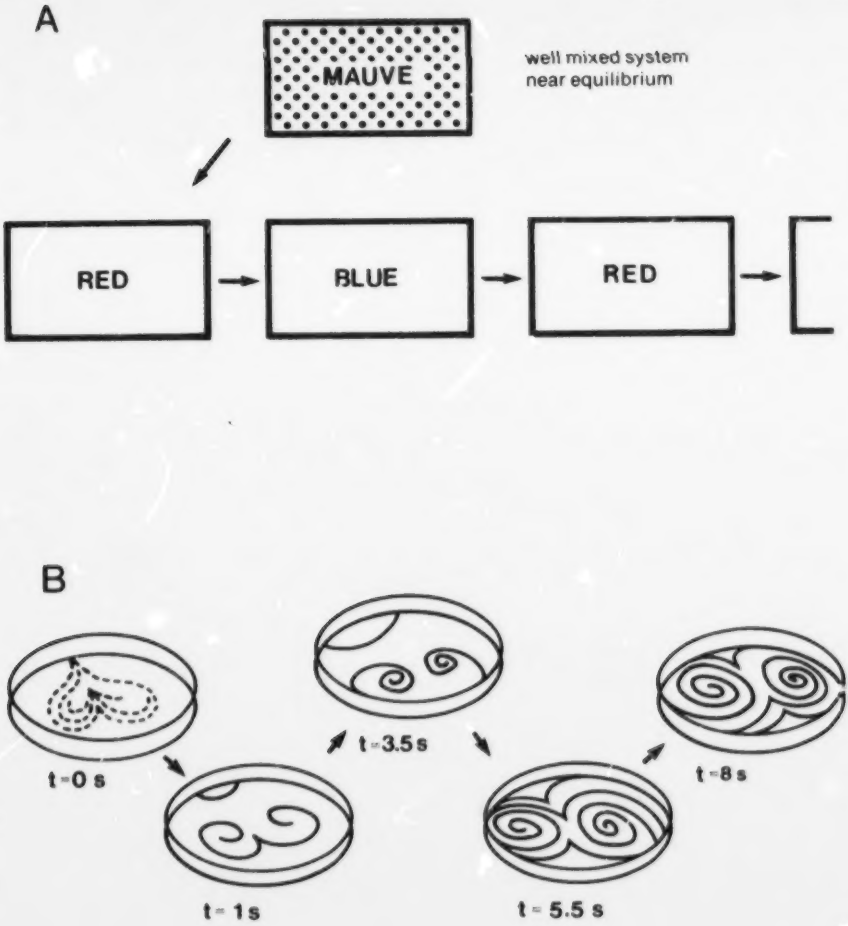


FIG. 5. (A) The entire system can begin to oscillate with a regular period from red to blue and back. (B) In an unstirred dish this is just one example of the multitude of patterns of spiral waves that can emerge spontaneously. (C) Many spatial and spatio-temporal patterns are possible, and in each case the diffusion fluxes, and the spatial organization are mutually dependent aspects of a single entity.

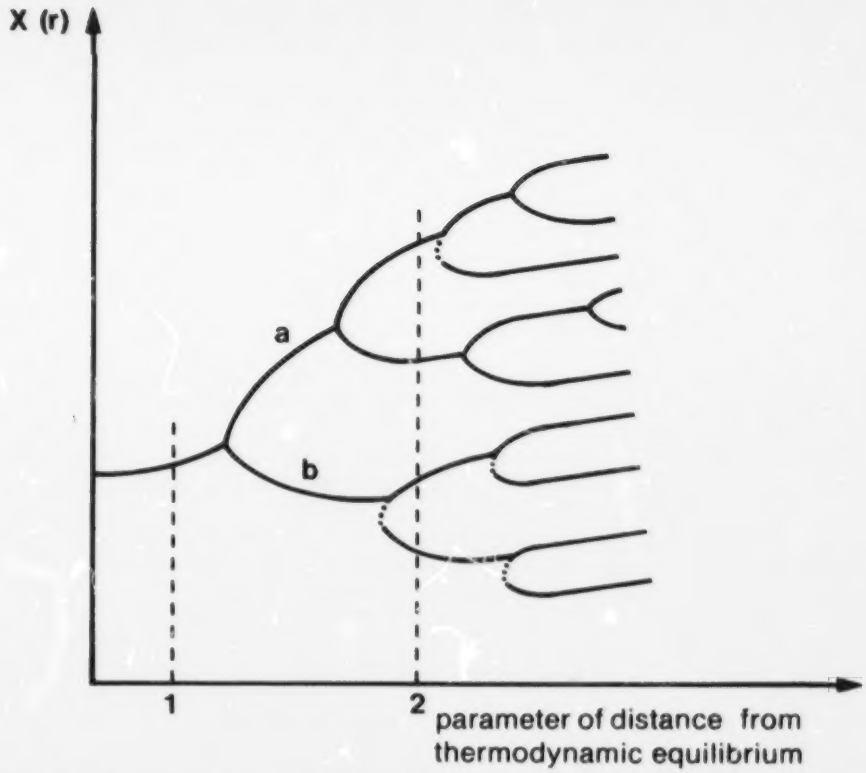


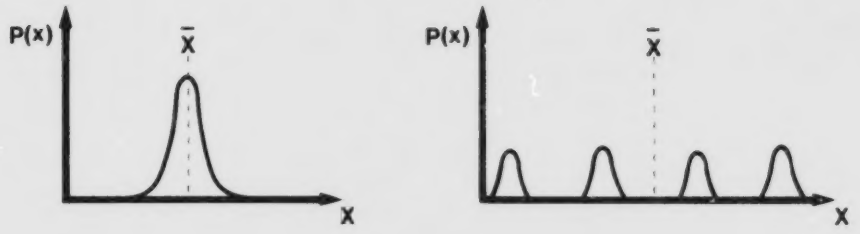
FIG. 6. As the system is driven further from thermodynamic equilibrium, suddenly instead of being characterized by a unique and trivial solution, there appears a bifurcating tree of possibilities. Branches correspond to solutions of different "symmetry" — they differ qualitatively.

What is new, and important, is that different solution branches can emerge which are qualitatively different from each other. We have, therefore, a *non-conservation* of symmetry, and hence of the number and nature of the "qualities" which characterize the system.

In one stride we have moved from the relative banality of simple arithmetic to the quantitative modelling of morphogenetic processes whereby structure and function emerge, where the qualitative differences of the living world appear, and in which we find creation instead of conservation.

We may understand more deeply the "mechanics" of what is happening by examining more closely what system sampling, and repeated experiments would "reveal" at points (1) and (2) of Fig. 6. In case (1) there is a single unique solution. In fact, the real system, made of billions of molecules, is fluctuating around this value, but with very small deviations. Hence, the mean value of X is a very good representation of what is occurring.

However, if we sample a single system at (2), then we shall find it in one of four possible branches at any given time. Occasionally, it may spontaneously jump to another branch, but we shall not find it at the "average" of all four possible solutions. Similarly, repeating



the experiment will show us simply that the system can be expected on one of the four branches. If we know that it took branch a at some earlier time, then we can assign a higher probability to finding the system in one of the upper two configurations.

This tells us that near to bifurcation points the macroscopic, reduced description in terms of kinetic equations of average densities, breaks down. The system becomes "sensitive" to hitherto unimportant internal or external fluctuations, which can drive the unstable system off to some new branch and to some new average description. At a bifurcation point the existence of small, possibly random, fluctuations can introduce new dimensions into the average description, and we observe that the symmetry of the solutions is broken as form is created.

Evolution represents a dialogue between the real, rich micro-detail of the system, and the simpler deterministic average behaviour which we have considered to be adequate to represent it. These interesting new perspectives can be illustrated by a simple, amusing example of morphogenesis — Origami, the Japanese art of paper folding.

Origami — the Emergence of Form

Origami consists in taking a flat sheet of paper and making folds in it in such a way as to produce a stable configuration which suggests to us some familiar object or creature. Consider, for example, some forms that can be made from a square sheet of paper (Kennaway 1980).

Figure 7 indicates the number of folds necessary to create the various forms and also the moments when "bifurcation" occurs and two objects become different. The various forms obtained by folding the paper are stable configurations in that the particular fold lines that have been impressed on it by the irreversible performance of work. They are self-consistent realizations of form and fold lines, otherwise the paper would either spring back or tear. Several important considerations emerge which are of great relevance to the understanding of evolutionary systems.

Firstly, folding the paper generates new traits and new images, and the various branches differ qualitatively because of this. Initially, we start with a piece of paper that has few attributes. It is square and white. After the folding process many attributes can be assigned to the various forms. Wings, legs, petals, volume, shape, and elegance have emerged, and the problem of choosing the "best" form, involves the comparison of objects spanning different dimensions, i.e., it is a "value judgement" and in evolution depends on whether or not these various features are compatible with the environment. What is more, these traits emerge at certain moments in the folding process, and each object has a past in which it was not what it is now, and a future in which it would cease to be what it is now, if the folding were to continue. Modelling which is based merely on what a system *is* now is descriptive, and does not contain the past nor the future of the system. As a tool to explore the future evolution of the system it must, therefore, fail at some point. What we require

in a model is that it should somehow *generate* the structure of today, from among other possiblities, and hence would be capable of examining the future without supposing that today's structure would continue forever.

Another important point arises when we open the paper and examine the pattern that the folds make. In some ways the pattern of folds represents the "DNA" of the object — the irreducible essence of the "seagull" or the "horse". The complexity of the object could be defined, for example, as being proportional to the number of folds required to make it, and the difference between objects as the number of folds separating them from the branch point of a common ancestor. However, all this overlooks an important point. The pattern of folds is in fact ambiguous, since it can be created in many ways, most of which do not lead to a recognizable object. Thus, the "essence" of the "seagull" or the "horse" is contained not only in the folds but also in the order in which the creases are made! The dynamics, or timing of the folding matters, and history plays a vital role in the nature of an object.

All that traditional, reductionist science has to say about the evolutionary tree of Fig. 7 is that all these objects weigh the same and are made of paper. Further tests would reveal with greater certainty that they were indeed made of paper, the type and quality of which could be identified.

However, this is all quite beside the point. What is important are the attributes that we assign to an object which distinguish it from other objects, and these traits are *not conserved*. It is the difference between things that are vital and those that are the basis of language itself. A dictionary which necessarily defines words in terms of other words, is nevertheless useful because, when coupled with experience, it enables the communication of the essential differences between things.

An interesting, if slightly absurd, idea would be to imagine that these paper forms needed to be renewed after a certain time. If paper were scarce, then they would have to

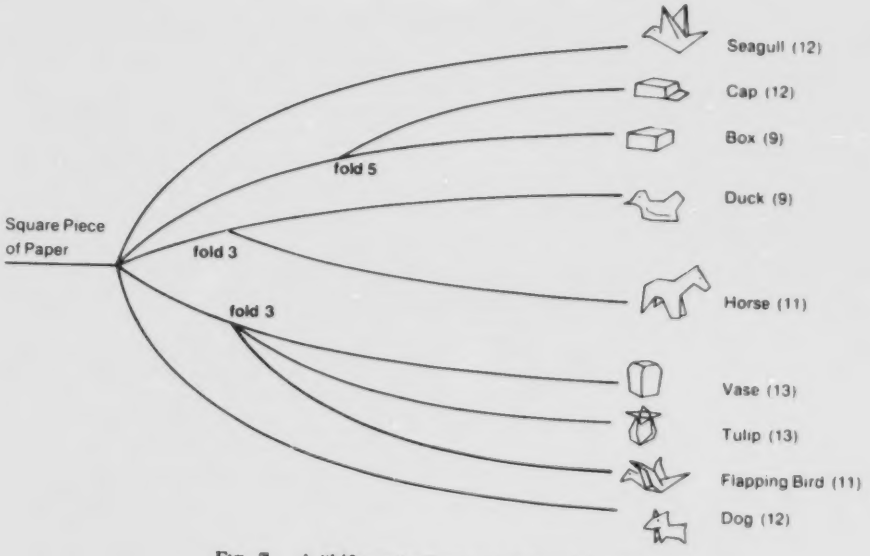


FIG. 7. A "bifurcating" tree of Origami forms.

merit their renewal, and this could be accorded if each form possessed at least some unique quality. If however, there were more than one form with the same set of qualities, or traits, then a competition must ensue resulting in only one "survivor". In this way the survival of a paper "duck" depends on the fact that the other forms are not duck-like, and this means of selection will give rise to an "origami world" of increasing diversity, populated by not-incompatible forms. Similarly, in the natural world, if the ability to tap some resource is associated with the possession of some particular trait or quality, then the same general picture of development would hold, and evolution in a given region would be characterized by a particular set of not-incompatible forms. From time to time, however, one could imagine that, as a result of gradual changes in particular species, or in the spectrum of available resources, the whole ecosystem might suffer a sudden re-organization. In some sense, then, the "unit of selection" is the entire ecosystem, which will produce a set of "mutually compatible" forms, probably with no particular global optimality.

Having somewhat belaboured the conceptual implications of these new ideas, let us turn now to some more practical equations concerning the understanding of real ecosystems.

Modelling Simple Ecosystems

The essential new idea in the foregoing discussion is that in order to "understand" a particular ecosystem we must comprehend how it became what it is. That is, we must look at its structure as representing the "accumulation" of successive adaptations to circumstances at different moments in the past. If an ecosystem persists in time, then we should not simply describe it, but rather attempt to establish the "reasons" for its stability.

We have seen that evolution is to be viewed as a "dialogue" between the simple reduced description of our minds (differential equations and average behaviours) and the real richness of the system, which is constantly producing fluctuations and abnormalities that probe the stability of the existing state. We have a dialogue between what is compressed into the simple model at a particular time and all that is not.

There are basically three different types of "fluctuations" which are present in the system: (a) density fluctuations of the variables (b) environmental fluctuations which appear as fluctuations in the parameters c) fluctuations in the internal structure of the elements – mutations or innovations.

Each of these classes deserves attention. In the first case, (a), the manifestly different spatio-temporal structures observed in ecosystems could result from self-organization processes such as those discussed in connection with the Brusselator. In marine systems patchiness, shoaling, the behaviours of predators and their prey, and the various vertical and horizontal structures could possibly be stable, inhomogeneous, and perhaps time dependent solutions of the governing non-linear dynamics. In fact, very little research has been done on this matter. The "reaction–diffusion" equation of chemistry has been studied widely, but the fact that organisms may diffuse through space differently than molecules (i.e., not necessarily Fickian) makes this issue a potentially fruitful field for research.

As far as environmental or parametric fluctuations, (b), the same applies. Studies of simple systems of equations clearly show that the effect of parametric fluctuations on the behaviour of non-linear systems is highly non-trivial. It has been shown, for example, that entirely new macroscopic system states can appear when a parameter fluctuates around its average value. Diminishing the noise, while maintaining the same average parameter value, results in a switch to a state that may be characterized as "quiet". The importance of environmental fluctuation has been commented on by May, who was able to show in a fairly convincing manner that "niche separation" in a simple one-level ecosystem would

be directly related to the degree of uncertainty in the environment. Since then, however, the whole field of stochastic non-linear differential equations has opened up, and the new advances should prove to be of great relevance to eco-science (e.g., Horsthemke and Lefever 1984).

The dialogue between the simple reduced model of an ecosystem and the third type of fluctuation, (c), mutations or innovations, is particularly interesting in that it gives rise to what is usually referred to as evolutionary ecology. Here I briefly summarize some of the results which seem to suggest another fruitful, but relatively unexploited avenue of research.

The dialogue between the macrostructure of an ecosystem and the "mutations" which occur in it is comprised of two distinct phases: (1) an initial stochastic period and (2) deterministic selection. Let us ignore phase (1) for the moment and discuss deterministic selection.

Consider as an example, the simplest possible ecosystem, a single species growing according to the logistic equation,

$$(5) \quad \frac{dx}{dt} = bx(1 - x/N) - mx.$$

This equation, describing the growth of a species x in a system with limited resources has a stable, stationary state, $x'' = N(1 - m/b)$.

Let us consider, however, the effect of the arrival in the system of a "mutant", x', that is different from x. For example, x' competes with x to an extent β for the limiting resource N. The mutant is characterized by some other birth rate b' and death rate m'. We shall suppose that after being subjected to some initially stochastic events, it has managed to survive and to become sufficiently numerous to be able to speak of a "density" (albeit very low) of mutants. The system equations become:

$$(6) \quad \begin{aligned} \frac{dx}{dt} &= bx(1 - (x + \beta x')/N) - mx \\ \frac{dx'}{dt} &= b'x'(1 - (x' + \beta x)/N') - m'x'. \end{aligned}$$

Now, we will not discuss whether x' is "better" or "worse" than x. Instead we shall ask whether or not x' can invade the system. This question is decided by testing the stability of the pre-existing state, $x'' = N(1 - m/b)$; $x' = 0$. If it is stable, then x' cannot invade the system. If it is unstable, invasion can proceed.

A simple stability analysis shows that the condition for x' to invade is,

$$(7) \quad N'(1 - m'/b') > \beta(N(1 - m/b)).$$

When this condition is fulfilled, x' will grow. Thus, if we supposed that a different mutant x' appeared with randomly scattered values of b', m' and N', then the deterministic equations would allow only those mutants to grow whose parameters satisfied condition (7). It would reject all the others.

Two cases arise. If the mutation x' were in total competition with x, then $\beta = 1$, and the condition becomes:

$$(8) \quad N'(1 - m'/b') > N(1 - m/b)$$

Hence, as a result of random mutations, evolution within a given "niche" can lead to increased "exploitation", or increasingly efficient use of the resources. The important point in this case is that, the condition that allows x' to grow also ensures that x must decrease and disappear, as portrayed in Fig. 8.

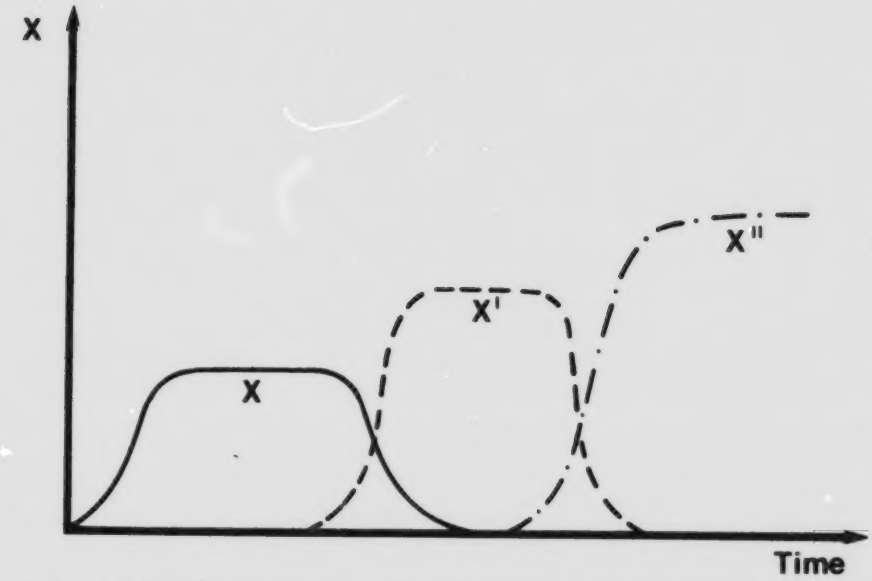


FIG. 8. Within the same "niche", we find successive replacement by more "effective" populations.

When overlap is not total, invasion is easier, since the value of $N'(1 - m'/b')$ need not be as high. What we shall observe, therefore, in a system with limited resources is that over a long time period an initially empty resource spectrum will gradually be filled by different populations, each adapted to a certain range of resources. Also, within any particular range or type of resource the efficiency of exploitation will increase irreversibly. This result can be extended to situations where genetics are explicitly considered, but these slightly more complicated equations do not lead to a different qualitative result. (Fig. 9).

Of course, these very impressive statements are almost tautological when applied to such a simple system — they correspond roughly to *proving* that, if a species is born more and dies less, then it will grow in the system. However, the idea might be interesting if it could be generalized to more complicated ecosystems, where the parameters represented the multiple effects of many interactions and mechanisms, in which case our method could be used to show under which circumstances certain types of evolution would be favoured.

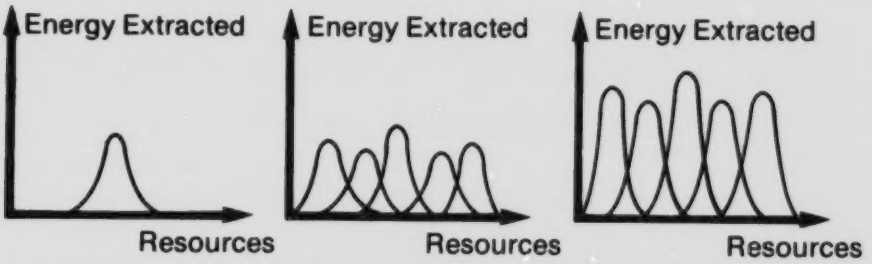


FIG. 9. Evolution will lead to filling an empty resource spectrum, and to increasing the efficiency of exploitation.

This generalization is indeed possible, Allen (1976). We suppose that n species are interacting according to some dynamic equations,

$$(9) \quad \frac{dx_i}{dt} = Gi(x_1, x_2 \ldots\ldots\ldots x_n).$$

Let us further assume that these n populations have attained a stable stationary state, $x_1^o, x_2^o, \ldots\ldots\ldots x_n^o$. If some new populations occur, for example, following the appearance of a new allele, then the stability matrix for the expanded system will be:

$$(10) \quad \left| \begin{array}{c|c} \begin{matrix} \frac{\partial G_1}{\partial x_1} - \lambda & \frac{\partial G_2}{\partial x_1} & \cdots\cdots\cdots & \frac{\partial G_n}{\partial x_1} \\ \frac{\partial G_1}{\partial x_2} & \frac{\partial G_2}{\partial x_2} - \lambda & & \vdots \\ \vdots & & \text{OLD SYSTEM} & \\ \frac{\partial G_1}{\partial x_n} & \frac{\partial G_2}{\partial x_n} & \cdots\cdots\cdots & \frac{\partial G_n}{\partial x_n} - \lambda \end{matrix} & \begin{matrix} \frac{\partial G_{n+1}}{\partial x_1} & \cdots\cdots & \frac{\partial G_{n+\Delta}}{\partial x_1} \\ & & \\ & & \\ \frac{\partial G_{n+1}}{\partial x_n} & \cdots\cdots & \frac{\partial G_{n+\Delta}}{\partial x_n} \end{matrix} \\ \hline \begin{matrix} \frac{\partial G_1}{\partial x_{n+1}} & \cdots\cdots\cdots & \frac{\partial G_n}{\partial x_{n+1}} \\ \vdots & & \vdots \\ \text{ZERO} & & \\ \vdots & & \vdots \\ \frac{\partial G_1}{\partial x_{n+\Delta}} & \cdots\cdots\cdots & \frac{\partial G_n}{\partial x_{n+\Delta}} \end{matrix} & \begin{matrix} \frac{\partial G_{n+1}}{\partial x_{n+1}} - \lambda & \cdots & \frac{\partial G_{n+\Delta}}{\partial x_{n+1}} \\ \vdots & \ddots & \vdots \\ \text{NEW PART} & & \\ \vdots & \ddots & \vdots \\ \frac{\partial G_{n+1}}{\partial x_{n+\Delta}} & \cdots & \frac{\partial G_{n+\Delta}}{\partial x_{n+\Delta}} - \lambda \end{matrix} \end{array} \right| = 0$$

Because of the zeros in the lower left hand corner expressing the fact that there is no *systematic* production of the $n + \Delta$ by the n (otherwise the $n + \Delta$ would have already been present) we find: stability matrix of whole = (stability matrix of old) × (stability matrix of new).

But we had supposed the old system had attained a stable stationary state, so that if an evolutionary step is to occur, then it can result only from the existence of a positive root of the stability matrix for the new populations — evaluated at the existing stationary state, $x_1^o, x_2^o, \ldots\ldots\ldots x_n^o$ and $x_{n+1}, x_{n+2}, \ldots = 0.$, that is, there exists a positive λ satisfying

$$(11) \quad \left| \frac{\partial G_{n+i}}{\partial x_{n+j}} - \lambda \right| = 0, \qquad i, j = 1, 2, \ldots \Delta,$$

at $x_1^o, x_2^o \ldots\ldots\ldots x_n^o$ and $x_{n+1}, x_{n+2} \ldots x_{n+\Delta} = 0.$

This general result has been applied to several different ecological systems (Allen 1975, 1980), but for the sake of brevity I will simply summarize the main conclusions.

If, for example, we consider a simple two-level predator–prey system, then we find that the occurrence of random mutations in the individuals of each level leads to an irreversible evolution. If the equations are:

$$\frac{dx}{dt} = bx(1 - x/N) - sxy \quad \text{and} \quad \frac{dy}{dt} = sxy - my,$$

then we obtain the stationary solution $x^o = m/s$ and $y^o = (b/s)(1 - m/Ns)$.

However, applying our general stability criterion to this system, we find that random mutations will lead to an increase in b and N and a decrease in m. The evolution of s will have no clear direction, corresponding as it does to an "arm race" between predator and prey. But the result of the varying parameters will be that the ratio of predator to prey will tend to increase, i.e.,

$$y^o/x^o = (b/m)(1 - m/Ns)$$

will rise over long times (Fig. 10).

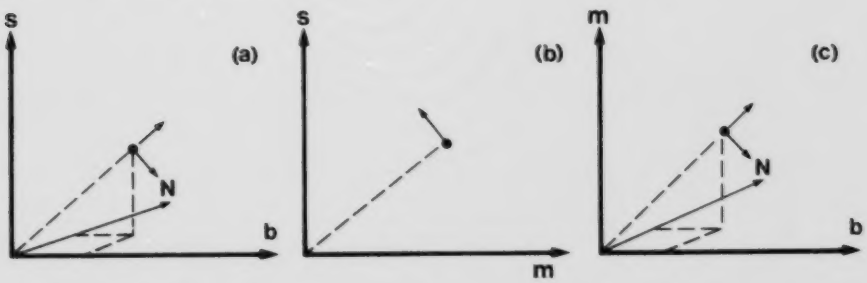


FIG. 10. The evolution of the prey (a), and of the predator (b) lead to increasing both b and N and decreasing m for the whole system (c). In turn the ratio of predator to prey numbers increases over long times.

In fact, the evolution of the two levels will tend to transfer biomass from the lower to the higher trophic level, and in some sense the evolution of the feeding efficiency of the "prey"; tends to serve the predator. In effect the prey is a "tool" through which the predator takes primary sustenance, and a more effective "tool" improves the predator's ability to feed. However, as the biomass becomes more concentrated in the higher level, so in their turn do these higher level species become potential "prey" for some new invader.

A similar study of the evolution of a parasite/host system showed that, in a stable environment, the co-evolution of both the parasite and its host would gradually progress from a very lethal parasitism to an innocuous presence, and possibly even to a symbiosis. However, new, lethal parasitic forms might still appear from time to time through the spontaneous generation of mutants.

Another interesting application of these methods has been to the evolution of "specialists" or "generalists" in ecosystems. This issue is of great interest because it gives rise to the community-level concept of the "connectivity" of ecosystems, and therefore to the questions which I posed at the beginning concerning the "reasons" why certain "arrows" exist between certain "boxes" but not others. Here I will treat a very simple case of a one-level system, but it nevertheless indicates the potential of the approach. Here, we return in a sense to the simple "logistic" model, but we treat more explicitly some of the processes which are included in the particular value, b, assigned to the "birth rate".

Consider a resource base of density c, in which we find a species x, that "extracts" E units of energy from each particle during a time τ. (This argument is borrowed partly from that of Maynard-Smith, but our criterion (11) states it in a more rigorous manner.) Each individual of the species x feeds on a certain band of resource type having a width w (a volume in hyperspace),

$$(13)\quad \frac{dx}{dt} = \frac{\alpha E w c}{1 + \alpha \tau w c}\, x(1 - x/N) - mx.$$

Using the evolutionary criterion (11), we see that N will increase, m will decrease and, as a result of random mutations, $\frac{\alpha E w c}{1 + \alpha \tau w c}$ will also increase.

However, different ways of increasing this latter term are possible, and according to the prevailing circumstances, some are more effective than others. For example, we must assume that there exists an inverse relationship between the amount of energy that can be extracted per particle, and the width of the resource band utilized by an individual. It is the non-linearity of these two factors which will be important. If the system is "rich" that is if $c \gg 1/\alpha\tau w$, then,

$$(14)\quad \frac{\alpha E w c}{1 + \alpha \tau w c} \to E/\tau$$

and the most effective amplification will occur for those mutations that increase the value of E/τ independently of the width utilized. As there is an inverse relation between the width and the value we can expect for E/τ, it follows that we shall observe an evolution towards specialization in these circumstances.

If, on the contrary, we are in a "poor" system, with $c \ll 1/\alpha\tau wc$, then because,

$$(15)\quad \alpha E w c/(1 + \alpha \tau w c) \to \alpha E w c,$$

evolution will tend to increase Ew, and lead to generalists exploiting a resource width that will depend on the precise form of the curve relating E and w.

This result can be extended to consider the morphological diversity that characterizes evolved ecosystems. If we suppose a resource base of length L and density c, then we may ask how many species we should expect to find sharing these resources as a set of "not-inconsistent" forms. As we have mentioned, May already showed that the separation between two species (the mythical Competitive Exclusion Principle) should be proportional to the amount of environmental fluctuation. However, this separation is expressed here in terms of the "width" occupied by the species, so it is possible to combine our results with those of May to obtain an expression for the expected morphological diversity (in a single level, simple, highly artificial example, of course). If the number of species is n, and their "niche" separation d, then we should find that,

$$d/w_s = \epsilon|\sigma^2| \text{ where } \sigma \text{ reflects environmental variability}$$

Since

$$n = L/d,$$

then

$$(16)\quad n = L/(\epsilon|\sigma^2|w_s),$$

or

$$(17)\quad nw_s = L/\epsilon|\sigma^2|.$$

But the width occupied by a species is given by the variability, v, of the species multiplied by the width occupied by individuals,

(18) $w_s = vw_i$

and w_i is inversely related to resource density c. Therefore, we may write,

(19) $nv = L/(\epsilon|\sigma^2|w_i) = Lc/\epsilon|\sigma^2|$,

which tells us that morphological diversity (e.g., of feeding apparatus) should be proportional to resource volume, Lc, and inversely proportional to the degree of environmental fluctuation.

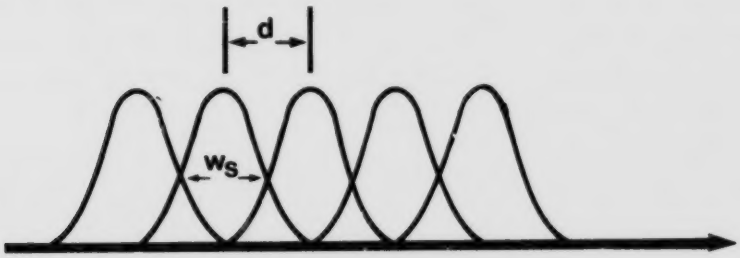


FIG. 11. A resource spectrum is supposed occupied by different populations, each of width w_s and separated by "distance" d.

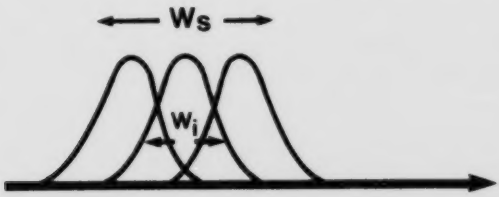


FIG. 12. The width occupied by a species is assumed to depend on that occupied by individuals, multiplied by some factor of variability.

Some partial confirmation of this relationship has been obtained. It concerns "Darwin's Finches" which inhabit the Galapagos Islands and which have been the subject of several careful investigations over the years. As is well known, the Islands are home to some 14 species of finch, which are generally not found elsewhere (Fig. 13). The "explanation" of their diversity is commonly assumed to be "ecological release", whereby the empty niches of the islands allowed the evolution of the few original species into the various possible varieties (Lack 1947). In Table 1 we see the numbers of species which occupy the different sized islands. One may already note some confirmation of the proportionality of number of species to resource volume in the tables, but this observation is too simplistic to carry much weight. If we consider the particular type of vegetation that each species occupies, then we can draw a more exact comparison. Bowman (1961) completed a careful

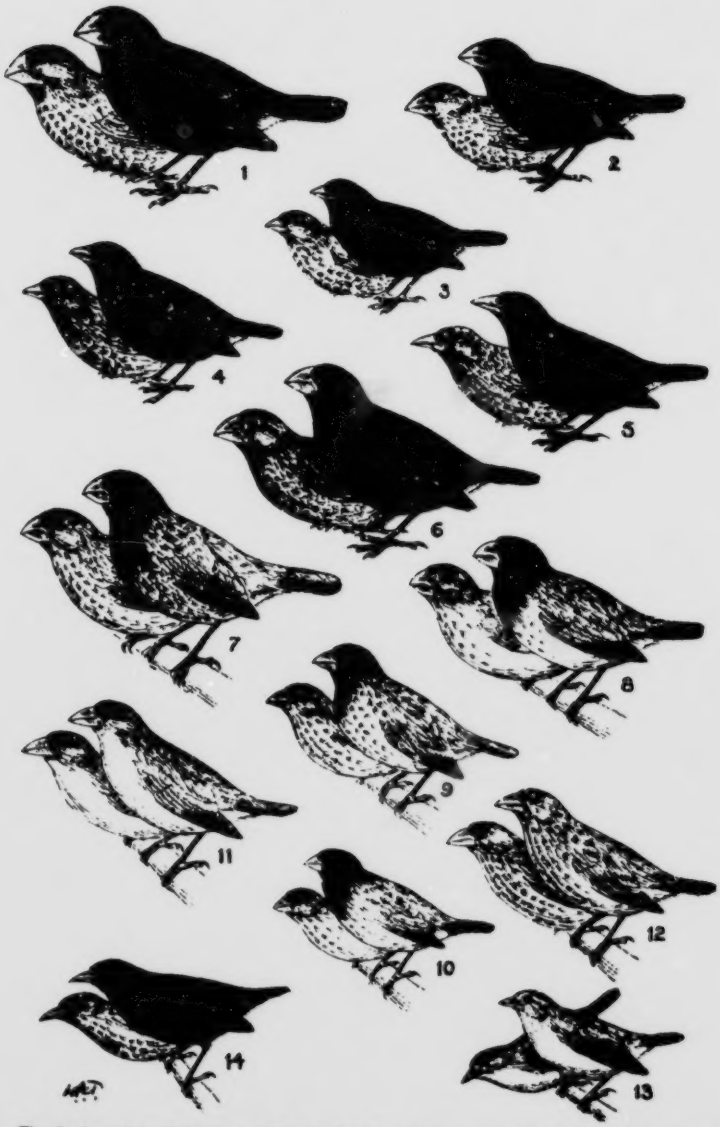


FIG. 13. The finches of the Galapagos. Male and female of each species. (1) *Geospiza magnirostris* Gould — large ground-finch, (2) *Geospiza fortis* Gould — medium ground-finch, (3) *Geospiza fuliginosa* Gould — small ground-finch, (4) *Geospiza difficilis* Sharpe — sharp-beaked ground-finch, (5) *Geospiza scandens* (Gould) — cactus ground-finch, (6) *Geospiza conirostris* Ridgway — large cactus ground-finch, (7) *Camarhynchus crassirostris* Gould — vegetarian tree-finch, (8) *Camarhynchus psittacula* Gould — large insectivorous tree-finch, (9) *Camarhynchus pauper* Ridgway — large insectivorous tree-finch on Charles, (10) *Camarhynchus parvulus* (Gould) — small insectivorous tree-finch, (11) *Camarhynchus pallidus* (Sclater and Salvin) — woodpecker-finch, (12) *Camarhynchus heliobates* (Snodgrass and Heller) — mangrove-finch, (13) *Certhidea olivacea* Gould — warbler-finch, (14) *Pinaroloxias inornata* (Gould) — cocos-finch. From *Darwin's Finches* Lack (1953), Cambridge University Press.

TABLE 1. Number of finch species occupying different sized islands of the Galapagos group.

Islands	Number of species
Large islands	
Albemarle	10
Indefatigable	10
James	10
Charles	9
Abingdon	9
Chatham	9
Intermediate islands	
Narborough	9
Bindloe	7
Small islands	
Wenman	5
Tower	4
Culpepper	4
Hood	3

TABLE 2. Relationship between morphological diversity of feeding apparatus and resource volume for the transitional vegetation found on all the islands. A "prediction" of total diversity seems possible but not the precise nature of the co-existence.

Island	Species	Resources	Total beak variety
Indefatigable	Magnirostris	Large seeds	21.6
Albemarle	Fortis	Medium seeds	20.6
Charles	Fuliginosa	Small seeds	18.9
James	Scandens	Cactus	18
Chatham	Fortis	Large seeds	17.6
	Fulaginosa	Medium and small seeds	
	Scandens	Cactus	
Tower	Magnirostris	Large seeds	13
	Difficilis	Small seeds	
	Conirostris	Cactus	
Narborough	Magnirostris	Large seeds	11.4
	Fortis	Medium seeds	
	Fuliginosa	Small seeds	
Hood	Conirostris	Large seeds and cactus	10.2
	Fuliginosa	Small seeds	
Culpepper	Conirostris	Large seeds	9.6
	Difficilis	Small seeds and cactus	
Wenman	Magnirostris	Large seeds	7.16
	Difficilis	Small seeds an cactus	

study of the transitional vegetation and of the finches which occupy it. He identified the diet of each type of finch, measured the beak sizes and variations, and proved that beak size was directly related to the diet. The results of this study are shown in Table 2, where the species occupying the transitional zones are shown, together with the manner in which they divide the resources, and the total diversity of beak measurements. One of the most interesting results is that the seed spectrum, which on large islands is divided between three specialists, is used by only two on Chatham. However, the measurements made by Bowman reveal that one of the two species is extremely variable, and in effect occupies two "niches".

Our formula (19) seems to approximate the results, and we see that such evolutionary arguments cannot predict precisely which "species" will occupy a given set of resources, nor exactly how these resources will be partitioned — circumstances which depend on the particular history of the system. However, that some relationship (like 19) exists between resource volume, environmental fluctuation, and morphological diversity implies that the resource base and the ecosystem connectivity are not arbitrarily related, and that some "predictive" statement may be possible.

Other interesting applications of criterion (11) have been made, for example to the study of "Dove/Hawk" strategies, and to the evolutions of the division of labour and of altruism. The method agrees somewhat with the "Evolutionary Stable Strategies" developed by Maynard-Smith (1979), but I believe that, despite the appeal of Maynard-Smith's beautiful, intuitive arguments, the real "dynamics" of such situations is complex and should be studied with rigor.

I have so far neglected the first phase of a mutation — the stochastic, early period when chance inevitably plays a large role. When a mutant first appears, it is necessarily a single individual. In the turmoil of the world, its survival through a given period of time is not assured by the average birth or death rates, which are applicable only to a population. Instead, the individual merely has a probability of surviving, of reproducing or of dying over the short term and these survival rates give rise to a stochastic equation that can be used to describe the probability of its extinction or survival over the long term. If the probability (per unit time) of reproduction (requiring perhaps a series of mechanisms) is A, and that of dying is D, then we can represent the probability of reproduction or mortality of a mutant by the form,

$$(20) \quad A'/D' = (A/D)(1 + \delta),$$

where δ is the fraction by which the mutant is "better" or "worse" at survival than the parent population. A simple calculation (Bartholomay 1958) reveals that the probability of survival through time t of a mutant population, starting initially from a single individual, is:

$$(21) \quad P_{\text{surv}}(t) = \frac{\delta}{1 + \delta - e^{-\delta(D'A/D)t}}$$

If we ask for the probability of a mutant surviving n generations, then the result is,

$$(22) \quad P_{\text{surv}}(n) = \frac{\delta}{1 + \delta - e^{n\delta/1+\delta}},$$

which is a general form, true for any mutant arriving in any simple ecosystem. In Fig. 14 we see the plots of $P_{\text{surv}}(n)$ for different values of n. Most importantly, we see that the probability of survival even for a much improved mutant (+10%) is only 10%, which means that the stochastic phase is a real barrier to the emergence of innovations. In particular, we find that the very sharp distinction in our deterministic analysis between "favourable" and "unfavourable" mutants (which translate into either "domination" or "extinction") is smoothed considerably in this more correct, stochastic picture. What it says is that the "power" of selection is limited, not infinite, and that for rather long times, a whole "spectrum" of relative performance could continue to exist in the system.

In order to understand better the implications of this result for the evolutionary process, let us consider the probability for the "production" of an effective mutant. Clearly, the overall rate of evolution of a species will depend on the rate at which different mutants are produced, combined with the rate at which the system selects the "more adapted" from among the population. Let us suppose that the probability of the production of a mutant

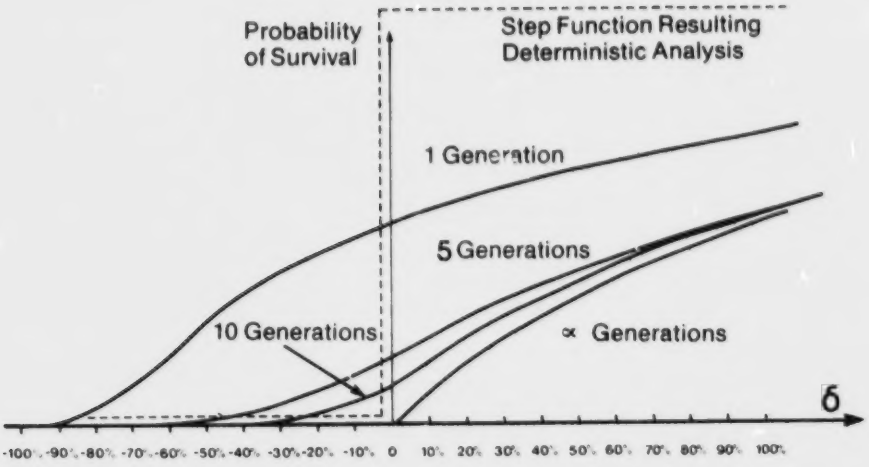

Probability of Survival
Step Function Resulting Deterministic Analysis
1 Generation
5 Generations
10 Generations
∝ Generations
δ
-100% -90% -80% -70% -60% -50% -40% -30% -20% -10% 0 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

FIG. 14. The form of the probability curves for the probability of surviving n generations.

with advantage δ is given by a Gaussian distribution as:

$$(23)\quad M(\delta) = \frac{2\pi}{\sigma} e^{-\delta^2/2\sigma^2}.$$

We can thus write the joint probability that a mutant with advantage will both appear, and survive n generations as,

$$(24)\quad P(\delta, n) = M(\delta) \cdot P_{sum}(n) = \frac{2\pi}{\sigma} e^{-\delta^2/2\sigma^2} \frac{\delta}{1 + \delta - e^{-n\delta/1+\delta}}.$$

This function is illustrated in Fig. 15 and will naturally vary according to the value of the variance δ. However, it should be clear that the relative probabilities of the appearance and survival of mutants that are slightly retrograde, neutral or positive, differ by surprisingly small amounts over at least 10 or 20 generations. As the number of specified generations is increased, the probability of survival decreases, and the curve moves to the right. Thus, the probability of survival for one generation is very nearly symmetric about the zero axis, but as n increases the curve progressively shifts to the right. We can calculate the "most frequent evolutionary event" that will persist for long times by calculating value of the δ where the greatest probability for survival occurs after an infinite number of generations. However, for the most part, the system will be characterized by a quasi-neutral drift with a really advantageous mutation occasionally sweeping deterministically through the system. If these occasional events are accompanied to any degree by assortative mating, they will "fix" neutral mutations in the surviving populations.

Another interesting point is that we can examine the relative probabilities of survival over long times of different, but equally advantageous, strategies for survival. For example, two populations may achieve the same value of A/D in two ways — either with birth and death rates that are both high, or with a combination of low birth and death rates. To survive over a long time t means that species with the former strategy mutates faster (more births) but must survive more generations to avoid extinction. Allen and Ebeling (1983) show that for any $\alpha > 1$,

$$(25)\quad \alpha P(\delta, \alpha n) > P(\delta, n),$$

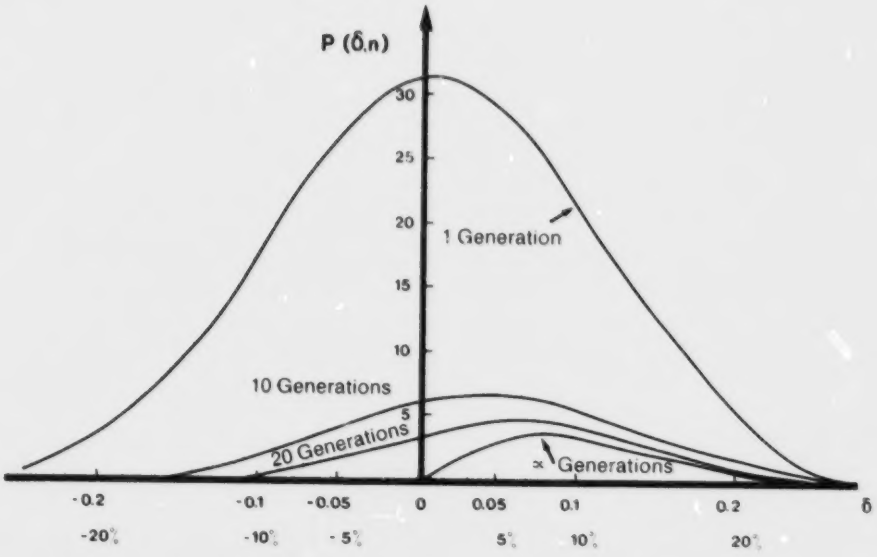


FIG. 15. Probability that a mutant with advantage will appear and survive n generations.

which implies that evolutionary steps come more frequently from the "high birth rate, high death rate" end of the species spectrum than from the other. This small effect will therefore tend to penalize, in the long run, an evolution towards species having extremely long organism life-times. This finding is in agreement with the idea advanced by Conrad (1979) that "adaptability" itself is a product of evolution. Another result which I shall briefly summarize is that stochastic considerations (which reduce the importance of selection forces) show us how a natural system can "experiment" with apparently disadvantageous mutations for a significantly long time. Thus, "evolution" is *not* purely a "hill-climbing" process, but instead allows the probing of the surrounding "landscape" and the location of other "hills". This result allows for the appearance of a division of labour, altruism or interdependence into a system. For, in a homogeneous system that obeys a strictly neo-Darwinian determinism, where even the first effects of an invading mutant must be advantageous for it to succeed, such things are impossible. Whereas the stochastic "softening" of selection, together with the possibility of spatial inhomogeneity allow initially disadvantageous mutants to persist for a while.

In fact, our analysis shows that the "unit of selection" can change, and some of the "fitness" of each individual can be contained in the spatial interactions of the system. Figures 16 and 17 show how an altruist population, BB, can invade a system, providing that diffusion and mixing are not too strong. Once this occurs, however, the system can be invaded by mutations, X_{AB}, which further reduce diffusion or increase cooperativity, and we detect a fundamental bifurcation in the behaviour of populations. Either the mutants pursue an individualist, type r strategy, reflecting strong mixing and turbulence in the environment, or they can behave "collectively" and require concerted, spatial structuring of the group. In the second case once a group exists, complex behavioural changes could be initiated through imitation and learning processes, and evolution will pass out of the "genetic" mode and into the realm of socio-cultural processes.

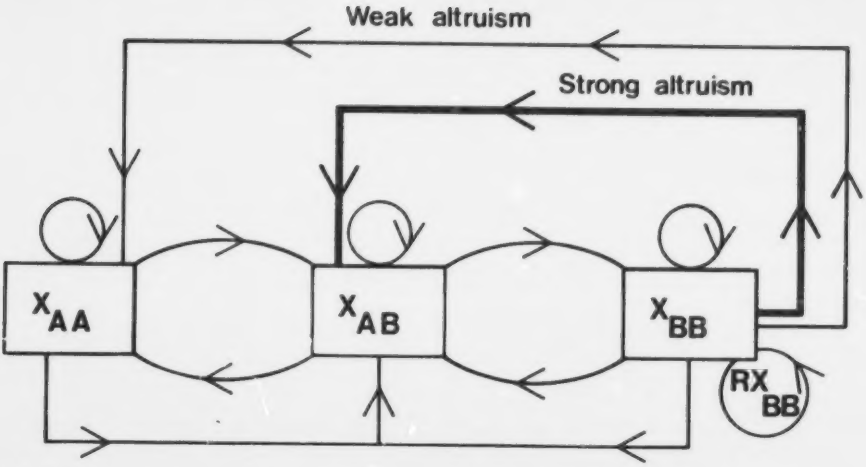


FIG. 16. When the act of altruism affects only a certain spatial area, then if the average separation of X_{AB} and X_{BB} is less than that of X_{AA} and X_{BB}, altruism can grow in the system.

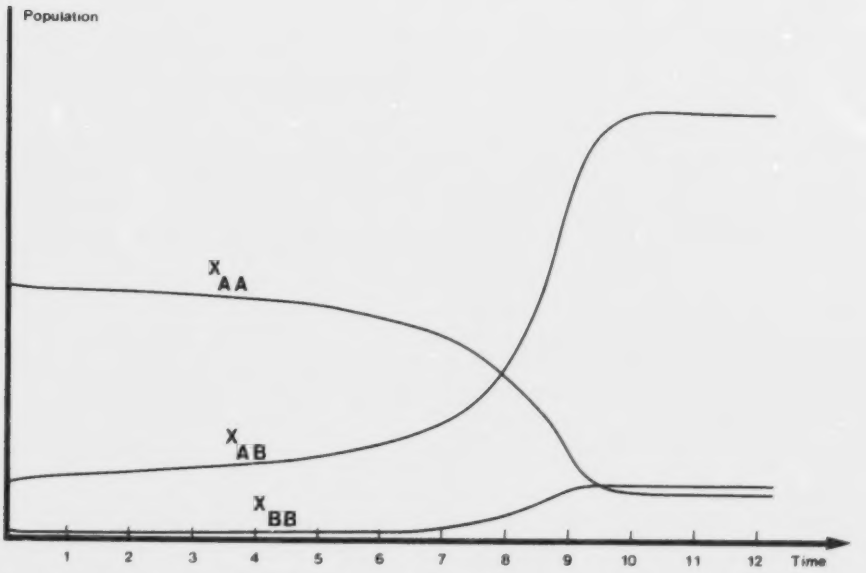


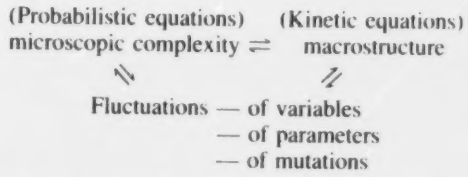
FIG. 17. Invasion of mutations X_{AB} and a fundamental bifurcation in the behaviour of populations.

Conclusions

The original context for this paper was the relation between thermodynamics and ecology. In fact, this relation turns out to be complex. It is, indeed, the study of thermodynamics which reveals that systems far from thermodynamic equilibrium can evolve

towards the creation of organization, diversity, and complexity. Thus, it is thermodynamics by itself that shows how in such systems there will be no direct relations between a few simple, macroscopic variables which will succeed in governing the system. Systems will be history dependent. Dissipative structures may offer a basic paradigm with which to understand the living world. On the other hand, they do not offer any immediate, easy answers which directly "explain" the peculiarities of particular ecosystems.

Instead, the theory of dissipative structures tells us that the link between microscopic and macroscopic behaviour, i.e., the source for the "explanation" of a reduced description of reality, is more complicated than it was for simple, or equilibrium systems. The evolution of far from equilibrium systems is generated, not by simple universal laws, but instead by a dialogue between microscopic reality and its reduced representation as a simple macroscopic structure, or as a "system dynamics" of differential equations, as depicted:



Rather than relying solely on descriptive modelling, each of the three types of "exposition" gives criteria for stability that would "explain" the observed structure. Therefore, a model of interacting populations should be studied in the context of a "space" larger than that which it occupies at a given instant. Thus, to understand properly a system, we should study the effects of density fluctuations in the variables in the context of a spatial disaggregation. Similarly, we should subject the system's spatial and relational structure to environmental and "mutational" fluctuations. Clearly, in the given examples the idea of "invadability" means that we study the stability of the population "types" in an ecosystem within the larger "space" of the different types that could exist. We do this because this is clearly something like the manner in which nature itself itself probes the world at each moment.

Of course, the work presented here offers only the first small step in examining this dialogue; but, hopefully, this hesitant step will be "amplified", and we shall obtain a deeper understanding of the world in which we live — a world which we now desecrate and exploit in almost total ignorance of the consequences.

References

ALLEN, P. M. 1975. Evolution in a predator prey ecology. Bull. Math. Biol. 37: 389–405.

1976. Evolution, population dynamics and stability. Proc. Nat. Acad. Sci. 73(3): 665–668.

1980. La modelisation des systemes complex. (Course notes available from the author.)

ALLEN, P. M., AND W. EBELING. 1983. Evolution and the stochastic description of simple ecosystems. Biosystems 16: 113–126.

BARTHOLOMAY, A. T. 1958. On the linear birth and death processes of biology as a Markov chain. Bull. Math. Biophys. 20(2): 97–119.

BOLTZMANN, L. 1872. Weitere Studien ueber das Waermegleichgewicht unter Gasmokuelen. Wien. Ber. 66: 275–370.

BOWMAN, R. I. 1961. Morphological differentiation and adaptation in the Galapagos finches. Charles Darwin Foundation for the Galapagos Islands — Contribution 1; University of California Publications in Zoology, Vol. 58.

CONRAD, M. 1979. Bootstrapping on the adaptive landscape. Biosystems 11(2 and 3): 167–182.

HORSTHEMKE, W., AND R. LEFEVER. 1984. Noise induced transitions. Springer Series in Synergetics, Berlin.

KENNAWAY, E. 1980. Origami — paper folding for fun. Octopus Books, London.

LACK, D. 1947. Darwin's finches. Cambridge Univer-

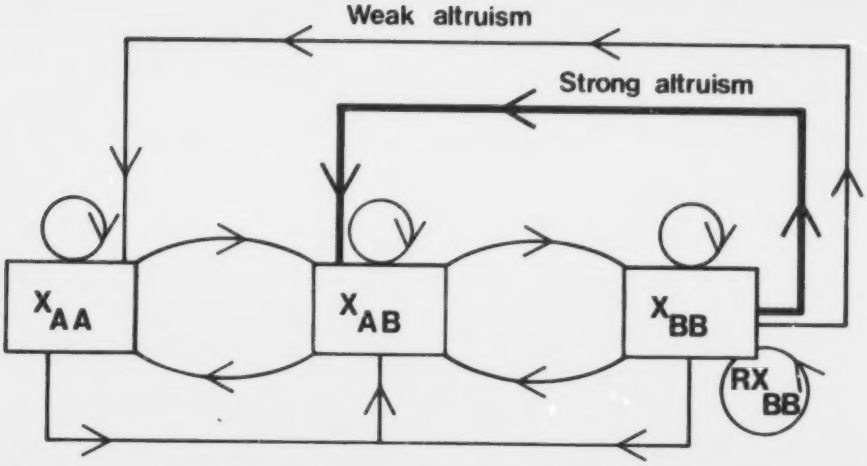


FIG. 16. When the act of altruism affects only a certain spatial area, then if the average separation of X_{AB} and X_{BB} is less than that of X_{AA} and X_{BB}, altruism can grow in the system.

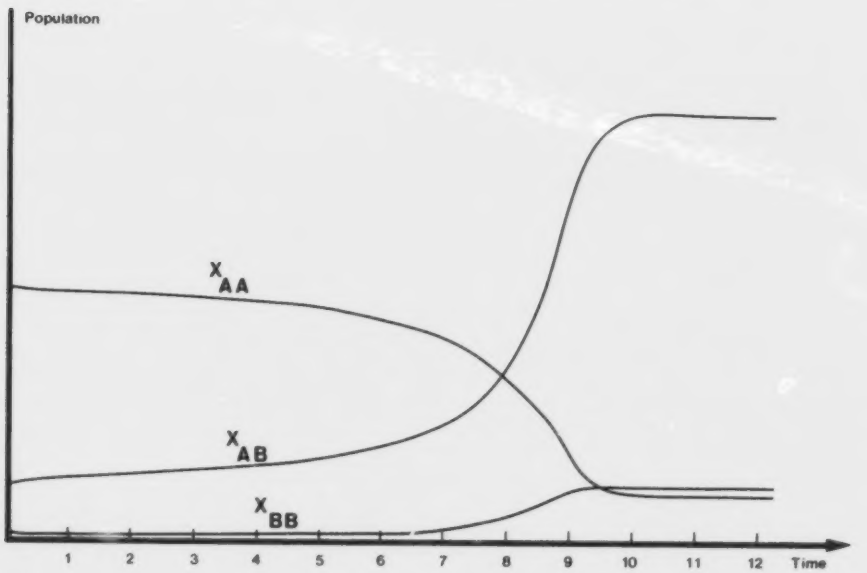


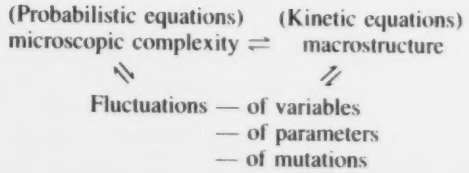
FIG. 17. Invasion of mutations X_{AB} and a fundamental bifurcation in the behaviour of populations.

Conclusions

The original context for this paper was the relation between thermodynamics and ecology. In fact, this relation turns out to be complex. It is, indeed, the study of thermodynamics which reveals that systems far from thermodynamic equilibrium can evolve

towards the creation of organization, diversity, and complexity. Thus, it is thermodynamics by itself that shows how in such systems there will be no direct relations between a few simple, macroscopic variables which will succeed in governing the system. Systems will be history dependent. Dissipative structures may offer a basic paradigm with which to understand the living world. On the other hand, they do not offer any immediate, easy answers which directly "explain" the peculiarities of particular ecosystems.

Instead, the theory of dissipative structures tells us that the link between microscopic and macroscopic behaviour, i.e., the source for the "explanation" of a reduced description of reality, is more complicated than it was for simple, or equilibrium systems. The evolution of far from equilibrium systems is generated, not by simple universal laws, but instead by a dialogue between microscopic reality and its reduced representation as a simple macroscopic structure, or as a "system dynamics" of differential equations, as depicted:



Rather than relying solely on descriptive modelling, each of the three types of "exposition" gives criteria for stability that would "explain" the observed structure. Therefore, a model of interacting populations should be studied in the context of a "space" larger than that which it occupies at a given instant. Thus, to understand properly a system, we should study the effects of density fluctuations in the variables in the context of a spatial disaggregation. Similarly, we should subject the system's spatial and relational structure to environmental and "mutational" fluctuations. Clearly, in the given examples the idea of "invadability" means that we study the stability of the population "types" in an ecosystem within the larger "space" of the different types that could exist. We do this because this is clearly something like the manner in which nature itself itself probes the world at each moment.

Of course, the work presented here offers only the first small step in examining this dialogue; but, hopefully, this hesitant step will be "amplified", and we shall obtain a deeper understanding of the world in which we live — a world which we now desecrate and exploit in almost total ignorance of the consequences.

References

ALLEN, P. M. 1975. Evolution in a predator prey ecology. Bull. Math. Biol. 37: 389–405.

1976. Evolution, population dynamics and stability. Proc. Nat. Acad. Sci. 73(3): 665–668.

1980. La modelisation des systèmes complex. (Course notes available from the author.)

ALLEN, P. M., AND W. EBELING. 1983. Evolution and the stochastic description of simple ecosystems. Biosystems 16: 113–126.

BARTHOLOMAY, A. T. 1958. On the linear birth and death processes of biology as a Markov chain. Bull. Math. Biophys. 20(2): 97–119.

BOLTZMANN, L. 1872. Weitere Studien ueber das Waermegleichgewicht unter Gasmokuelen. Wien. Ber. 66: 275–370.

BOWMAN, R. I. 1961. Morphological differentiation and adaptation in the Galapagos finches. Charles Darwin Foundation for the Galapagos Islands — Contribution 1; University of California Publications in Zoology, Vol. 58.

CONRAD, M. 1979. Bootstrapping on the adaptive landscape. Biosystems 11(2 and 3): 167–182.

HORSTHEMKE, W., AND R. LEFEVER. 1984. Noise induced transitions. Springer Series in Synergetics, Berlin.

KENNAWAY, E. 1980. Origami — paper folding for fun. Octopus Books, London.

LACK, D. 1947. Darwin's finches. Cambridge Univer-

sity Press; 1953. Sci. Am. 188(4): 66–71.

MAYNARD-SMITH, J. 1979. Game theory and the evolution of behaviour. Proc. R. Soc. London B. 205: 475–488.

NICOLIS, G., AND I. PRIGOGINE. 1977. Self-organization in non-equilibrium systems. Wiley Interscience, New York, NY.

PLATT, T., K. H. MANN, AND R. E. ULANOWICZ. 1981. UNESCO Press, Paris. 156 p.

PRIGOGINE, I. 1947. Études thermodynamique des phénomènes irréversibles. Desoer, Liége.

PRIGOGINE, I., AND P. M. ALLEN. 1982. The challenge of complexity. *In* Schieve and Allen [ed.] Self-organization and dissipative structures. Texas University Press, Austin, TX.

Thermodynamics and Growth of Organisms in Ecosystems

A. I. Zotin

Institute of Developmental Biology, Academy of Sciences of the USSR, Moscow, USSR

Introduction

There is a long history of attempts to apply thermodynamics to the description of processes in living systems. Despite these appreciable efforts, the prognosis for developing a thermodynamics of biological processes remains doubtful. The difficulty lies in the fact that thermodynamics, as a general phenomenological theory of processes and phenomena occurring in nature, is far from complete. While two major branches of thermodynamics (classical thermodynamics and the thermodynamics of linear irreversible processes) have been elaborated, the phenomenological theory of non-linear irreversible processes is far from being perfected.

Thus, there are as yet no theoretical grounds for the systematic introduction of thermodynamic concepts and correlations into biology; although attempts in this direction are well-known, and I believe they have been useful. The approach of using thermodynamics to describe biological phenomena has been applied towards the construction of theories in ecology and biological oceanography. I am not a specialist in the field of ecosystems theory and cannot judge whether such attempts have been successful. However, for many years I have been studying the quantitative aspects of organism growth and have tried to construct a thermodynamic theory of ontogeny. My experience in developing this thermodynamic approach may also be of help in creating a similar theory for ecosystems. One should also not forget that the growth of organisms, and especially the energetic aspects or that growth are of great importance to the understanding of constituent processes occurring in ecosystems.

Now, concerning the quantitative theory of growth: At present a large number of growth equations have been delineated (Brody 1945; Richards 1959, 1969; Zotina and Zotin 1973; Zotin 1974; Walter and Lamprecht 1978; Żelawski and Lech 1980; Majkowski and Uchmański 1980; Parks 1982). Some of them, such as the Malthus and Verhulst equations of the Gompertz function, date back to the last century (Sandland 1983), but the others appeared later. By now there are 20–30 known equations that more or less adequately describe the growth of animals and plants. Obviously, some effort should be made to put the different theories of growth in order. Such work has been, is being and will be carried out until such time as a unified theory of growth has been elaborated.

The Types of Growth Equations

The unification of the different theories of growth has reached such a stage that three irreducible groups of equations can be distinguished from the whole suite of growth equations. The first group (Pütter–Bertalanffy) includes the variations on the most popular (especially in hydrobiology) Bertalanffy (1957, 1960) equation, which is expressed as,

$$(1) \quad \frac{dW}{dt} = NW^{m/n} - K_g W,$$

where W is the body weight; t, the time; N, m/n, and K_g are constants. By simple transformations (1) can be reduced to:

$$(2) \quad \frac{1}{W}\frac{dW}{dt} = K_s\left(\frac{W_m^b}{W^b} - 1\right),$$

where W_m is the weight of the adult animal; and b is a constant from the equation:

$$(3) \quad \dot{q}_{O_2} = aW^{-b}$$

In (3) $\dot{q}_{O_2}$ represents the intensity of oxygen consumption, and a is a constant. The well known Richards equation:

$$(4) \quad \frac{1}{W}\frac{dW}{dt} = \frac{K}{1-m}\left[\left(\frac{A}{W}\right)^{1-m} - 1\right],$$

and the Gompertz function:

$$(5) \quad \frac{1}{W}\frac{dW}{dt} = K_s e^{-at},$$

fall into this first group. The Richards formula is widely applied to describe plant growth (Richards 1969; Majkowski and Uchmański 1980), while the Gompertz function is often used when investigating the growth of mammals, birds (Laird et al. 1965; Walter and Lamprecht 1978; Parks 1982), and tumors (Laird 1964; Emanuel 1977). It is rather evident that the Richards equation is identical to the Bertalanffy expression (2). The matter is more complicated with the Gompertz function, however, it turns out that the Gompertz expression can be obtained from both the Bertalanffy (Zotin 1974) and the Richards (Sztencel and Lech 1980) equations. Let us rewrite the Bertalanffy equation (2) as:

$$(6) \quad \frac{1}{W}\frac{dW}{dt} = K^1(W^{-b} - W_m^{-b}).$$

Note that the constant b in the Bertalanffy equation was taken from equation (3). Substituting (3) into (6) we obtain:

$$(7) \quad \frac{1}{W}\frac{dW}{dt} = K_1(\dot{q}_{O_2} - \dot{q}_{O_2 m}),$$

where $\dot{q}_{O_2 m}$ is the respiration intensity of an organism reaching its limiting weight; and $K_1 = K^1 a$.

Brody (1945) showed that the change in the intensity of respiration during the growth of mammals and birds is described by the equation:

$$(8) \quad \dot{q}_{O_2} = \dot{q}_{O_2 m} + K_2 e^{-K_3 t}$$

Substituting this expression into (7), we obtain the Gompertz function in the form:

$$\frac{1}{W}\frac{dW}{dt} = K_4 e^{-K_3 t},$$

where $K_4 = K_1 K_2$.

It is easy to show that many other growth equations, for example Brody's exponential growth expression (Zotin 1974) and Schmalhausen's equation of parabolic growth (Winberg 1966), fall into the first group of growth formulae. Thus, the Pütter–Bertalanffy group includes most of the growth equations now in use.

The second group of equations not reducible to the Pütter–Bertalanffy group consists of those expressions for growth in which the limiting value is expressed not by a maximal weight, but rather by the time when the maximal weight is achieved.

The first such equation was proposed by us (Zotina and Zotin 1967, 1972) as:

$$(9) \quad \frac{1}{W}\frac{dW}{dt} = K_d(t_m - t).$$

Equation (9) adequately describes the growth of fishes (Zotina and Zotin 1967) and of some invertebrates (Knight et al. 1976; Klimenko 1971). A similar formula was advanced by Żelawski and Lech (1980) to describe plant growth:

$$(10) \quad \frac{1}{W}\frac{dW}{dt} = K_d\left[1 - \left(\frac{t}{t_m}\right)^{K_a}\right].$$

It is obvious that when $K_a = 1$, (10) is transformed into (9). Żelawski and Lech (1979) showed that the constant K_a can take on values in the range from 0 to 1, depending on the conditions under which the plant is growing. Thus, environmental factors influence the plant growth curve mainly by changing the value of the constant K_a.

Finally, the third independent group of growth expressions is related to the Robertson equation (Robertson 1908, 1923; Walter and Lamprecht 1978), which has the form:

$$(11) \quad \frac{1}{W}\frac{dW}{dt} = K_f(W_m - W).$$

This group includes the Verhulst equation (Walter and Lamprecht 1978; Sandland 1983):

$$(12) \quad \frac{1}{N}\frac{dN}{dt} = \beta\,\frac{1}{R}[R - N(t)],$$

where N represents the number of cells in an organism, and the Pearl logistic function (Walter and Lamprecht 1978):

$$(13) \quad N(t) = \frac{\beta N_0 e^{\beta t}}{\beta + \gamma N_0(e^{\beta t} - 1)}.$$

Since (13) is a solution of (12), the Verhulst and logistic functions fall into the Robertson group. This same class of formulae includes the Naydenov function (Walter and Lamprecht 1978) and the equation of Gines et al. (1974):

$$(14) \quad \frac{1}{W}\frac{dW}{dt} = K_f(W_m^b - W^b),$$

where b is the constant in (3). Setting $b = 1$ in (14) transforms it into the Robertson equation (11).

Thus, we have now identified three independent groups of growth equations. The next stage in describing the general process of organism involves establishing a framework of relations for unifying these three groups into one mathematical statement. This can be accomplished through a generalized thermodynamic theory of growth (Zotin 1974; Zotina and Zotin 1978) or by other, more formal methods (Turner et al. 1976; Presnov 1977; Savinov et al. 1977).

The Thermodynamical Approach

In the thermodynamics of linear irreversible processes the interrelationship between different processes occurring simultaneously in a thermodynamic system is described as:

$$(15) \quad \bar{I}_i = \sum_{j=1}^{m} L_{ij}X_j \; (i = 1, \ldots, n),$$

where $\bar{I}_i$ is the ith specific thermodynamic flow; X_j the thermodynamic force corresponding to the jth flow and L_{ij} are the linear phenomenological coefficients relating the forces and flows.

In the most general sense the processes involved in the growth and development of organisms consist of three types of phenomena: changes in weight (growth), the appearance of differences between various parts of the system (differentiation), and changes in the form and structure of the organism (pattern formation). Assuming that (15) is applicable to organism growth and designating the specific "flows" of growth by $\bar{I}_g = 1/W\, dW/dt$, differentiation by $\bar{I}_d = 1/W\, dD/dt$ and pattern formation by $\bar{I}_f = 1/W\, d,F/dt$, the interrelationship between growth, differentiation and pattern formation can then be written in general form as:

$$(16)\quad \frac{1}{W}\frac{dW}{dt} = K_{gg}X_g + K_{gd}X_d + K_{gf}X_f$$

$$\frac{1}{W}\frac{dD}{dt} = K_{dg}X_g + K_{dd}X_d + K_{df}X_f$$

$$\frac{1}{W}\frac{dF}{dt} = K_{fg}X_g + K_{fd}X_d + K_{ff}X_f,$$

where X_g represent the forces responsible for growth; X_d the forces behind differentiation; X_f the forces causing pattern formation; and K_{ij} are phenomenological coefficients.

In the thermodynamics of irreversible processes the forces appearing in the phenomenological equations of the type (15) are usually determined by empirical means. For example, Fick's law is invoked when analyzing diffusion processes, Fourier's law when investigating heat transfer, etc. If several irreversible processes proceed simultaneously in the system, their interrelationship is described by (15).

A similar approach can be taken in an effort to ascertain the concrete significance of forces in (16). Empirical laws, in this case the equations of growth, can be employed. To decide which of the numerous growth equations should be taken as the phenomenological laws of developmental biology one should consider the requirements imposed on the forces by the thermodynamics of irreversible processes. The principal condition limiting the choice of flows and forces is the requirement that they fulfill the relationship:

$$(17)\quad \psi = \frac{T}{V}\frac{d_iS}{dt} = \sum_{j=1}^{n} \bar{I}_j X_j,$$

where ψ is the specific dissipation function for the system; d_iS/dt, the rate of entropy production; T the absolute temperature; and V the volume (or weight of the system). Another, less significant constraint is that the variables determining the force must be of an intensive nature and their values must create some gradient. It is obvious also that the expressions modelling different forces must not be identical.

Merging Theory and Phenomenology

As mentioned above, three irreducible groups of equations may be identified out of the whole mass of growth expressions. These are the Pütter–Bertalanffy, the Zotina–Zotin, and the Robertson types of equations. It would be tempting to suggest that these three groups correspond to the three types of forces involved in the growth and development of organisms: growth, differentiation and pattern formation. The problem remains as to which equation should be taken as the phenomenological law of growth and which ones as the laws of differentiation and pattern formation.

There are some good reasons for choosing the Bertalanffy equation as the phenomenological expression for growth and the Zotina–Zotin equation for differentiation (Zotin 1974; Zotina and Zotin 1973, 1978):

1) The Bertalanffy equation is a better descriptor of the growth of animals than is the Zotina–Zotin relation (Walter and Lamprecht 1978).

2) when the Zotina–Zotin expression is chosen to represent differentation, the dimensionality $[MT^{-1}]$ is plausible because the specific protein responsible for differentiation is determined by its rates of symthesis and degradation;

3) The solution to the Zotina–Zotin equation is a Gaussian curve, which aptly describes the changes in differentiation during the growth and development of an organism — at first there are relatively few differentiations, then an increasing number and finally a decrease during the later stages of ontogenesis;

4) The combination of the Bertalanffy (2) and Zotina–Zotin (9) equations as in (16) gives the growth expression:

$$(18)\quad \frac{1}{W}\frac{dW}{dt} = L_{gg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{gd}(t_m - t),$$

which is capable of describing weight changes during the all life course of any animal. None of the other known growth relations is as comprehensive in description (Zotina and Zotin 1973; Zotin 1974).

Thus we assume that the Bertalanffy equation corresponds to growth and the Zotina–Zotin relation to differentiation. By exclusion the Robertson function in the form of (14) is taken here to be the phenomenological equation of pattern formation. This choice might be supported by the dimensionalities inherent in (16). For this purpose we combine (16) and (18) to obtain:

$$(19)\quad \begin{aligned}
\frac{1}{W}\frac{dW}{dt} &= L_{gg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{gd}(t_m - t) + L_{gf}(X_f^1)\\
\frac{1}{W}\frac{dD}{dt} &= L_{dg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{dd}(t_m - t) + L_{df}(X_f^1)\\
\frac{1}{W}\frac{dF}{dt} &= L_{fg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{fd}(t_m - t) + L_{ff}(X_f^1),
\end{aligned}$$

or, as dimensional equations:

$$(20)\quad \begin{aligned}
T^{-1} &= [L_{gg}] + [L_{gd}]T + [L_{gf}][X_f^1]\\
M^{-1}T^{-1}[D] &= [L_{dg}] + [L_{dd}]T + [L_{df}][X_f^1]\\
M^{-1}T^{-1}[F] &= [L_{fg}] + [L_{fd}]T + [L_{ff}][X_f^1].
\end{aligned}$$

where $X_f = \Gamma_f X_f^1$ and Γ_f are dimensional constants.

Unfortunately, we cannot obtain the dimensions of the force causing pattern formation from (20). Therefore, a more or less plausible proposal about the dimensionality of organism form should be advanced. Since changes in organism form are usually connected with the alteration of its surface, it is reasonable to assume that $[F] = L^2$. If so, it then follows from (20) that $[L_{fg}] = L^2M^{-1}T^{-1}$, and as $L_{fg} = L_{gf}$:

$$(21)\quad [X_f^1] = L^{-2}M$$

The dimensional relation (21) is consistent with the following expression for the force:

(22) $X_f = \Gamma_f\left(\frac{W_m}{\Omega_m} - \frac{W}{\Omega}\right),$

where Ω is the surface area of the organism and Ω_m the surface area of the organism after it has reached the stationary weight. As $\Omega = \alpha^1 V^k$ (where V is the body volume), α^1 and k are constants and $W = \rho V$ (ρ — body density), then $\Omega = \alpha W^k$ ($\alpha = \alpha^1/\rho$). The constant k in this expression is related to the parameter b in (3) in the manner $b = 1 - k$. Taking this into account (22) can be rewritten as:

(23) $X_f = K_f(W_m^b - W^b),$

and we obtain the Robertson equation in the form (14):

$$\frac{1}{W}\frac{dW}{dt} = K_f(W_m^b - W^b).$$

Now, substituting (23) into (19), we obtain:

(24)
$$\frac{1}{W}\frac{dW}{dt} = L_{gg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{gd}(t_m - t) + L_{gf}(W_m^b - W^b)$$

$$\frac{1}{W}\frac{dD}{dt} = L_{dg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{dd}(t_m - t) + L_{df}(W_m^b - W^b)$$

$$\frac{1}{W}\frac{dF}{dt} = L_{fg}\left(\frac{W_m^b}{W^b} - 1\right) + L_{fd}(t_m - t) + L_{ff}(W_m^b - W^b)$$

The first equation in the system (24) is the phenomenological growth equation.

Thus, the thermodynamic approach (albeit with the introduction of some hypothetical reasoning) allows for the unification of all the growth equations into a single description. It also results in a system of differential equations which covers all the basic phenomenological processes related to the growth and development of organisms.

Testing the Adequacy of the Description

The resultant phenomenological portrayal of growth describes the weight changes during the ontogenesis of many animals better than any other known equations insofar as it reproduces not only the weight increase during growth, but the stationary phase and weight decrease during senesence as well. Data on the growth of crickets, cockroaches, axolotls, hens, and white rats are shown in Fig. 1–5. In all the cases the curves obtained by fitting them with system (24) agreed well with actual observations.

The equations (24) were obtained under the assumption that the growth of organisms could be described by means of linear phenomenological equations (15). In thermodynamical terms this corresponds to the assumption that the weight changes in the system occur never far from equilibrium or stationary states. Although the growth and development of organisms progresses through a number of stationary states, it remains to be determined empirically whether linear relationships can be used to describe growth. It turns out that (24) is not applicable in all cases. In particular, it cannot be applied to describe the changes in the weight of humans throughout their lifetimes (Grudnitzky et al. 1971; Zotin 1974). Consequently, in some cases non-linear (in relation to forces) phenomenological equations of growth become necessary. Such a non-linear expression was obtained (Zotina and Zotin 1978) and used to describe changes in weight over the human lifespan (Konoplev et al. 1978) and to determine the maximum possible human lifespan (Prokofiev et al. 1982).

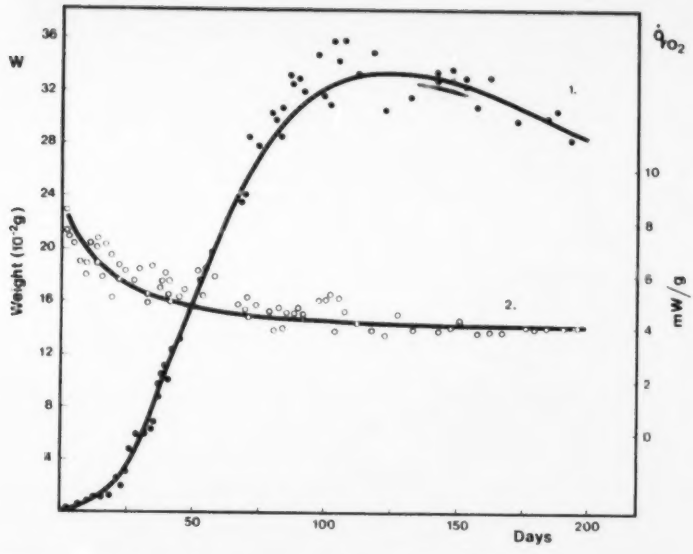


FIG. 1. Experimental data on the changes in weight (1) and respiration intensity (2) during the life of crickets when compared with the curves calculated using equations (24) and (26), respectively (Zotina et al. 1982).

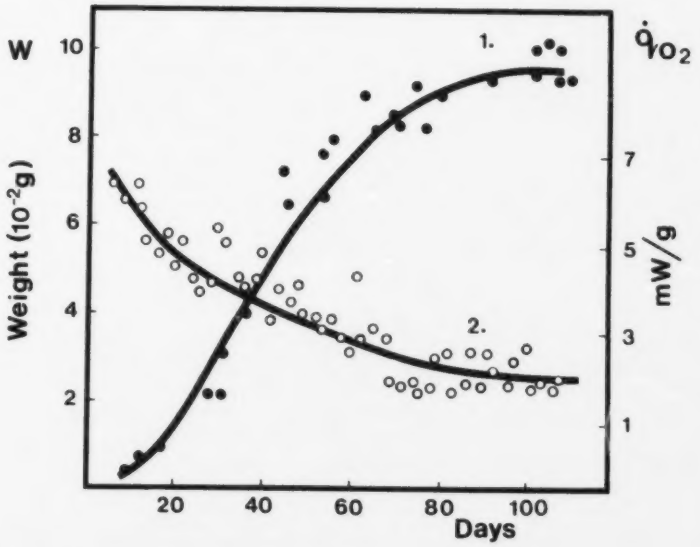


FIG. 2. Experimental data on the changes in weight (1) and respiration intensity (2) during the life of cockroaches when compared with the curves calculated using equations (24) and (26), respectively (Zotina et al. 1982).

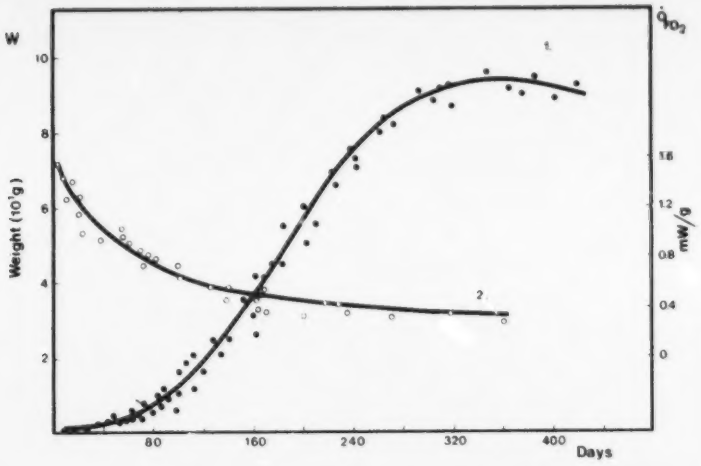


FIG. 3. Experimental data on the changes in weight (1) and respiration intensity (2) during the life of axolotls when compared with the curves calculated using equations (24) and (26), respectively (Zotina et al. 1982).

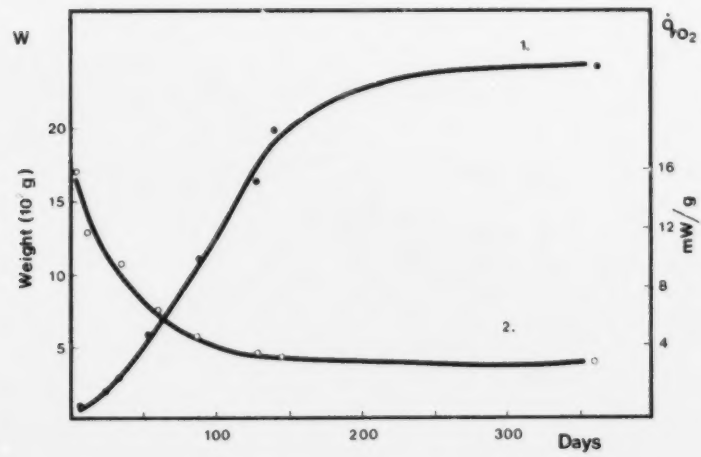


FIG. 4. Experimental data on the changes in weight (1) and respiration intensity (2) during the life of hens when compared with the curves calculated using equations (24) and (26), respectively (Zotina et al. 1982).

The agreement between (24) and the experimental data does not prove that the choice of flows and forces in this description is correct. As mentioned above, they must also obey condition (17). However, this requirement also provides a way to use thermodynamics not only to check the correctness of phenomenological growth equations, but also to obtain

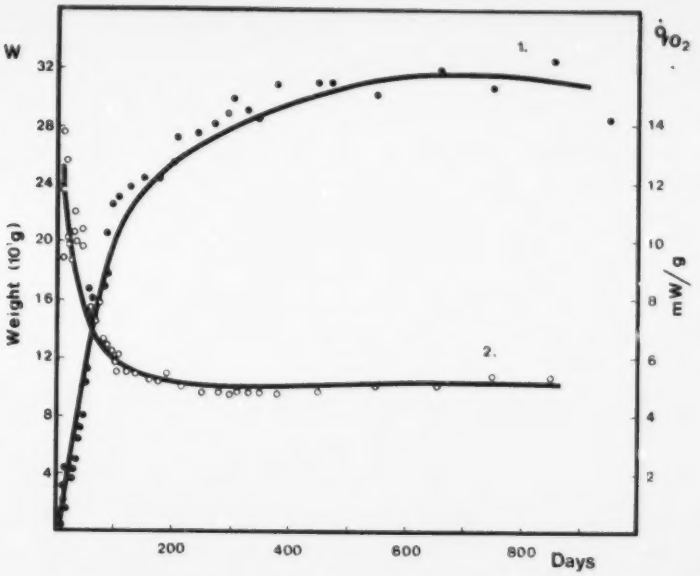


FIG. 5. Experimental data on the changes in weight (1) and respiration intensity (2) during the life of white rats when compared with the curves calculated using equations (24) and (26), respectively (Zotina et al. 1982).

an expression describing the changes in energetic metabolism during the course of growth and development.

In the final analysis all the dissipative processes in the living organisms are determined by respiration and glycolisis, which provide the main supply of energy to living systems. Therefore, it is appropriate to write (17) as:

$$\Psi = \dot{q}_{O_2} + \dot{q}_{gl} = \sum_{j=1}^{n} \bar{I}_j X_j.$$

Taking into account the fact that organism glycolysis under aerobic conditions is usually negligible, this simplifies to,

$$(25) \quad \dot{q}_{O_2} = \sum_{j=1}^{n} \bar{I}_j X_j.$$

Substituting the expressions for the flows and forces from system (24) into (25) we obtain:

$$(26) \quad \dot{q}_{O_2} = \Gamma_g \left[L_{gg} \left(\frac{W_m^b}{W^b} - 1 \right) + L_{gd}(t_m - t) + L_{gf}(W_m^b - W^b) \right] \left(\frac{W_m^b}{W^b} - 1 \right) +$$
$$+ \Gamma_d \left[L_{dg} \left(\frac{W_m^b}{W^b} - 1 \right) + L_{dd}(t_m - t) + L_{df}(W_m^b - W^b) \right] (t_m - t)$$
$$+ \Gamma_f \left[L_{fg} \left(\frac{W_m^b}{W^b} - 1 \right) + L_{fd}(t_m - t) + L_{ff}(W_m^b - W^b) \right] (W_m^b - W^b),$$

where Γ_g, Γ_d, and Γ_f are dimensional constants.

There are too many constants and parameters in this equation. However, the number of fitted coefficients can be significantly decreased by obtaining them from independent experimental data. In particular, the constant b can be calculated from observations on the relationship respiration and weight in the organisms under study. The coefficients L_{gg}, L_{gd}, L_{gf} are obtained by fitting (24) to the weight changes observed over the course of an organism's life. It follows from Onsager's reciprocal relations that $L_{dg} = L_{gd}$, $L_{gf} = L_{fg}$, $L_{df} = L_{fd}$. We also may assume that $\Gamma_g = \Gamma_d = \Gamma_f$. The values for the remaining coefficients are selected so as to make (26) describe the changes in respiration intensity over the course of the organisms is growth.

As a check to see how (26) could describe changes in energetic metabolism, we compared these predictions with the data on insects, amphibians, birds, and mammals shown in Fig. 1–5. It proved possible to describe these processes to a satisfactory degree.

Conclusions

Although the equations derived above for use in developmental biology are rather complicated, their applications should present no difficulties given the present state of computing techniques. They are evidence that the thermodynamics of non-equilibrium processes can be used to deduce quantitative relationships and equations of importance to developmental biology. The possibility is not to be excluded that thermodynamics might also be used to obtain phenomenological equations appropriate to the theory of ecosystems.

References

Bertalanffy, L. von. 1957. Quantitative laws in metabolism and growth. Q. Rev. Biol. 32(3): 217–231.

1960. Principles and theory of growth. p. 137–159. *In* Fundamental aspects of normal and malignant growth. Elsevier, Amsterdam.

Brody, S. 1945. Bioenergetics and growth. Reinhold, New York, NY.

Emanuel, N. M. 1977. Kinetics of experimental tumour processes. Nauka, Moscow. (In Russian)

Grudnitzky, V. A., A. I. Zotin, N. V. Terentieva, and N. Sh. Shagimordanov. 1971. Growth equation with allowances made for interaction of growth and differentiation. Ontogenesis 2(5): 451–454. (In Russian)

Klimenko, V. V. 1971. Equation of growth for *Bombyx mori* larvae. Ontogenesis 2(6): 617–625. (In Russian)

Knight, A. W., M. A. Simmons, and C. S. Simmons. 1976. A phenomenological approach to the growth of the winter stonefly, *Taeniopteryx nivalis* (Fitch) (Plecoptera: Taeniopterygidae). Growth 40(4): 343–367.

Konopleva, V. A., A. I. Zotin, E. A. Prokofiev, and N. Sh. Shagimordanov. 1978. Computer analysis of non-linear growth equations. p. 135–141. *In* Thermodynamics of biological processes. de Gruyter, Berlin.

Laird, A. K. 1964. Dynamics of tumour growth. Br. J. Cancer 18(3): 490–502.

Laird, A. K., S. A. Tyler, and A. D. Barton. 1965. Dynamics of normal growth. Growth 29(3): 233–248.

Majkowski, J., and J. Uchmanski. 1980. Theoretical foundations of individual growth equations in animals. Polish Ecol. Studies 6(1): 7–31.

Parks, J. R. 1982. A theory of feeding and growth of animals. Springer, Berlin.

Pressnov, E. V. 1977. Transformation time in the theory of growth. Qualitative investigation. Ontogenesis 8(4): 420–423. (In Russian)

Prokofiev, E. A., R. S. Zotina, and A. I. Zotin. 1982. Phenomenological equations of growth and its application for determine of maximal life-span. p. 56–66. *In* Mathematical developmental biology. Nauka, Moscow. (In Russian)

Richards, F. J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10: 290–300.

1969. The quantitative analysis of growth. p. 3–76. *In* Plant physiology. Academic Press, New York, NY.

Robertson, T. B. 1908. On the normal rate of growth of an individual and its biochemical significance. Wilhelm Roux' Arch. 25(3): s. 581–614.

1923. The chemical basis of growth and senescence. Lippincott, Philadelphia, PA.

Sandland, R. L. 1983. Mathematics and the growth of organisms — some historical impressions. Math. Sci. 8(1): 11–30.

SAVINOV, I. P., B. P. VASILIEV, AND V. M. SCHMIDT. 1977. One class of growth curves of plant. J. Gen. Biol. 38(3): 432–439. (In Russian)

SZTENCEL, I., AND A. LECH. 1980. On the convergence of the comprehensive (Richards) growth function to the Gompertz function. Acta Physiol. Plantarum 2(4): 319–321.

TURNER, M. E., E. L. BRADLEY, K. A. KIRK, AND K. M. PRUITT. 1976. A theory of growth. Math. Biosci. 29(3/4): 367–373.

WALTER, R., AND I. LAMPRECHT. 1978. Modern theories concerning the growth equations, p. 143–162. *In* Thermodynamics of biological processes. de Gruyter, Berlin.

WINBERG, G. G. 1966. The growth rate and metabolism intensity in animals. Adv. Modern Biol. 61(2): 274–293. (In Russian)

ZELAWSKI, W., AND A. LECH. 1979. Growth function characterizing dry matter accumulation of plants. Bull. Acad. Polon. Sci., Ser. Biol. 27(8): 675–681.

1980. Logistic growth functions and their applicability for characterizing dry matter accumulation in plants. Acta Physiol. Plantarum 2(2): 187–194.

ZOTIN, A. I. 1974. Thermodynamical approach to the problems of development, growth and aging. Nauka, Moscow. (In Russian)

ZOTINA, R. S., AND A. I. ZOTIN. 1967. Quantitative relationship between the size, age and fertility in animals. J. Gen. Biol. 28(1): 82–92. (In Russian)

1972. Toward a phenomenological theory of growth. J. Theor. Biol. 35(2): 213–225.

1973. United growth equations. J. Gen. Biol. 34(4): 606–616. (In Russian)

1978. Differential equations of developmental biology, p. 121–134. *In* Thermodynamics of Biological Processes. de Gruyter, Berlin.

ZOTINA, R. S., A. I. ZOTIN, AND E. A. PROKOFIEV. 1982. Phenomenological equations of constitutive processes, p. 169–177. *In* Mathematical developmental biology. Nauka, Moscow. (In Russian)

II. STATISTICAL MECHANICS

Statistical Mechanics in Biology — Applications to Ecology

L. E. H. Trainor

Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7

Introduction

Thermodynamics is a macroscopic approach to physical systems which emphasizes the interplay between dynamic and thermal effects. The basic concepts in the theory are anchored in the description of systems in equilibrium for which the entropy assumes its maximum value. For systems close to equilibrium, one can reasonably define thermal and dynamic fields (such as temperature and pressure) in which the gradients are gentle, and cause and effects are linearly related, e.g. through the so-called transport coefficients: temperature gradients give rise to heat flow, density gradients to diffusive flow, etc. In this near equilibrium regime, the so-called thermodynamic branch (Glansdorff and Prigogine 1971), the state of the system is uniquely described and one has such important results as the symmetry of the Onsager coefficients (relating various "forces" and "flows") and the fluctuation–dissipation theorem which states that equilibrium fluctuations are governed by the same phenomena that guide an isolated system back to equilibrium via dissipative forces. On the thermodynamic branch the second law can be supplemented by a principle of minimum entropy production consistent with constraints on the system.

Beyond the thermodynamic branch the behaviour of the system becomes much more complicated. The thermodynamic system eventually develops an instability as it is pushed farther from equilibrium, the solutions to the thermodynamic equations "bifurcate" and a new stability develops associated with stronger dissipation and a less symmetric (more patterned) state. As the system is pushed still farther from equilibrium, new bifurcations develop, eventually the concept of local thermodynamic equilibrium is lost, and chaotic or turbulent behaviour ensues. The Prigogine school has popularized the possible connection first pointed out by Turing (1952), between such "dissipative structures" which occur before turbulence sets in and pattern development and morphogenesis in biology, since biological systems are in many respects examples of systems operating in or between stable states far from thermodynamic equilibrium.

Statistical mechanics on the other hand is a microscopic approach to the macroscopic behaviour of physical systems. In the usual physics formulation the dynamics of the microscopic systems ("particles") are assumed known and statistical methods are employed in one way or another to carry out microscopic averages leading to a description in terms of macroscopic quantities associated with macroscopic behaviour. Statistical averaging is a process either for glossing over unknown information, (a somewhat dangerous but nonetheless common occupation), or for reducing the quantity of known information to manageable terms. In the former case some form of randomness or ergodic behaviour is usually assumed; in the latter case, some justifiable principles may be used to do the averaging sensibly. In microscopic $\hbar$-physics, for example, the averaging for systems of identical Bose–Einstein particles must be carried out in a different way from systems of identical Fermi–Dirac particles, at least at low temperatures.

It is well to keep in mind that the second law applies only to the low-grade or thermal energy in a physical–chemical system. For example, the famous Carnot cycle and its analysis as the most efficient possible heat engine is concerned with what fraction of *thermal* energy flowing between a high temperature and a low temperature reservoir may

be used to do useful work. Such considerations are not necessarily sufficient for studying cellular processes where high-grade energy, e.g. in an ATP molecule, may be transduced into other relatively high grade forms by the action of highly specialized molecular machines; these processes are not thermal processes and ergodic theorems are not applicable to them. An important principle seems to be that *self-organizing systems keep state availability selective* rather than random. This principle in no wise contradicts the second law, but supplements it. In colloquial terms one might paraphrase the situation for self-organizing systems by saying that "there is more out there than the second law". For one thing there are the physical laws which lead to structures as simple as the hydrogen atom and as complicated as a large enzyme. The second law pertains only to the thermal aspects of these interacting structures. Moreover, one can even argue that structures as complicated as enzymes are "smart structures" in their capacity to discriminate; Goodwin (1976) has suggested that such structures could be said to possess cognitive properties at some elemental level.

In most physical problems the level of such cognition is not very great, since the structures are very simple; a further simplification arises from the vast numbers of identical particles, which makes averaging relatively straightforward. In biological systems where the structures are complex, and where the complexity can grow and integrate enormously with the size of the system, the connection between structural complexity and degree of cognition becomes especially interesting, e.g. in physical theories of brain function. This connection may also be of special interest in ecology, where structure extends beyond the level of individuals and species to the system as a whole, with associated global levels of cognitive processes. (See remarks of Lewis Thomas, e.g. in *The Lives of a Cell*, concerning bee or termite colonies (1975)).

Use of Statistical Mechanics and Thermodynamics in Biology

We can distinguish two types of application of the concepts of statistical mechanics (SM) and thermodynamics (TD) to biological problems:

1) "Real" (i.e. physical) SM and TD
2) Metaphorical SM and TD.

In (1), the biological system in whole or in part is regarded as a physical system, however complex, and one can raise the question to what extent do the methods of SM and TD form a useful or predictive description. We shall argue below that while TD is often useful in biological description, statistical mechanics is seldom such. In (2) above, one attempts formal analogues between TD and SM schemes as employed in physics and some quite different schemes of interest in biology or ecology. For example, in population biology the "unit systems" may be individual animals of a species, in rough analogy to the particles of a gas or fluid in physics; and one may be interested in averaging and wish to draw analogy with techniques used in physical SM; but the problem does not in any direct way involve physical SM or TD. Generally, the attempts, originally by Kerner (1972), later by Trubatch and Franco (1974) and Lumsden and Trainor (1979a, 1979b, 1980), to found a statistical mechanics appropriate to biological problems fall into class (2) above, i.e., the class of metaphorical SM. These attempts while interesting in their own right, have succeeded only in part; e.g. the work of Lumsden and Trainor (1979a, 1980) showed how one could formally set the first order dynamics typical of chemical and biological rate equations into "Hamiltonian form", thus facilitating a comparison with the canonical formalism of conventional SM in physics; particularly interesting in this work is the representation of such systems as phase points moving on Dirac constraint hypersurfaces (Dirac 1950) in the analogue space to the phase space of SM in physics. The formal

presentation of time-dependent dynamics in terms of phase cloud compression and expansion was also achieved. However, as pointed out in a detailed analysis by Lumsden and Trainor (1984), the biological metaphor lacks several ingredients possessed by the canonical formalism of conventional statistical mechanics which renders a tight analogy difficult if not impossible.

Of particular note in physical SM is the concept that equal volumes of accessible phase space (i.e. equal hyperareas on the energy surface) have the same a priori probability; the quantum analogue is that every accessible state of a system has an equal a priori probability with every other such state. These ergodic or quasi-ergodic principles are crucial to a simple averaging process (e.g. replacing time by phase space averages); otherwise one has to deal with the complex problem of weighted regions of the phase space. Moreover, the dominant role played by the Hamiltonian in physical SM is based on the principle that the energy is the only *additive* constant of the motion (Landau and Lifschitz 1958). In non-Hamiltonian systems it is less clear which constants of the motion are additive and in what sense, particularly if the system dynamics are non-linear. Progress with metaphorical biological analogues to conventional formulations of statistical mechanics depends upon both interpretative and formalistic resolutions of these difficulties. In Volterra systems, for example, a constant of the motion can be identified only after a transformation from natural variables (such as animal populations or chemical concentrations) to unnatural variables (essentially exponentials of the natural variables). This transformation also converts Volterra systems from ones in which phase clouds (the phase fluids) are compressible in the space of natural variables, but incompressible in the space of unnatural variables, (Lumsden and Trainor 1979b). But since the whole exercise is to find a statistical mechanical metaphor in terms of natural variables, this limited example (Volterra systems) is not entirely encouraging for the development of a profound and useful metaphorical statistical mechanics in biology.

In short, while developments in non-linear dynamics and in the TD of systems far from equilibrium have been exciting, and even encouraging for useful biological description, one cannot claim great successes in biology for microscopic theory, that is for the statistical mechanical underpinnings of systems of either the physical or the metaphorical types.

Statistical Mechanics in Ecology

The problems which plague the development of a biological statistical mechanics also apply at the ecological level of description, but are even more troubling and extensive. Ecological systems manifest a complicated hierarchical structure which adds to the problem and to the challenge of developing a layered approach to statistical mechanics (see Lumsden and Trainor 1984). The essence of a simple and effective statistical mechanics, as stated above, is the ability to average sensibly over large amounts of microscopic information in order to reduce the description of the system to a few essential macroscopic variables. In the example of an ocean ecology one can ask the following questions preliminary to any progress with metaphors to physical SM:

— what are the particles in MSM (metaphorical SM)?

— what properties of the macroscopic system are to be deduced in order to compare with actual measurements?

— what averages are to be made and what are the principles which correspond to equal a priori probability of states and Hamiltonian dominance in theories of canonical ensembles in the physical paradigm?

— what are the relevant system dynamics and how do these bear upon statistical averaging?
— is there an equilibrium analogue? and are there analogues of steady states far from equilibrium?
— what are the pattern determining mechanisms? and are these primarily "dissipative" or primarily "structural"?

One cannot blithely use statistical mechanical considerations, whether for physical or metaphorical systems without painful justification. The development of a statistical mechanics in the much simpler context of physical systems at and near equilibrium has taken many decades to develop to its present artistic and useful but incomplete state. It would not be surprising, given the greater complexity of biological problems, and even more so, of ecological problems, if several decades passed again before a useful and artistic reduction of this problem has been achieved.

Structure–Function Relationships in Hierarchical Systems

A distinct even dominant feature of large ecosystems is their hierarchical organization, which can be viewed separately in structural and functional (or behavioural) terms and then related. At least in principle this can be done; in practice it is difficult. In physical analogues, structure relates at the microscopic level to what kinds of particles (quarks, atoms, molecules, etc.) are involved and what kinds of interactions (force fields) take place amongst them. The term "function" is generally not used in this context in physics because it has distinct mathematical connotations which are already overworked, but a roughly equivalent term is "behaviour"; e.g. one speaks about the behaviour of magnetic spin systems or of an excited gas of molecules. In some degree, the physicists' viewpoint is a cause and effect viewpoint, where particle properties and particle interactions are causes of (i.e. lead to) certain effects or behaviours.

The traditions in biology are somewhat different, where function (behaviour) has a purposeful or teleological connotation arising, perhaps, out of the context that biological survival is something more than mere physical existence. Nonetheless, physics does provide some useful and analyzable analogues to structure–function relationships in biology, and we will pursue a couple of examples in what follows.

First, however, a few remarks on the nature of ecological hierarchies are in order. These hierarchies are so complex that it is not at all clear how to define their structural elements nor how to assess interactions amongst them, however defined. In all probability there is no unique way of carrying out such a program of definition and assessment, which does not gainsay that each of several ways of approaching the problem might serve some limited utility and add a dimension of understanding to the totality of behaviour. In addition, there is the concomitant difficulty of identifying links between actual field measurements in experimental studies and quantities identified with behaviour in the theoretical modelling.

The difficulties in such a program are illustrated in several physical models based on hierarchical structures of hypothetical physical origin (Totafurno et al. 1980; Totafurno and Trainor 1982. These papers are hereinafter designated TLT and TT, respectively). In both cases we consider hierarchically organized systems, with structure at different levels, and the system as a whole in thermal equilibrium with a large heat bath. Structure in these systems is to be regarded as "actual" physical structure, and function is taken to be thermodynamic behaviour — sometimes in the system as a whole, sometimes apparent thermodynamic behaviour as expressed in terms of a reduced set of "level variables".

Two results emerge in these examples, which we first state and subsequently elaborate. The first result is established in TLT (1980) and the second in TT (1982):

Result 1

Autonomy of behaviour (i.e. function) at any particular level of a hierarchically structured system is enhanced if the strength of coupling between levels is graded with a tendency toward stronger coupling at lower levels.

Result 2

Ambiguities arise in modelling a given system as a hierarchically structured system; the same behaviour can be achieved in structural schemes which differ from each other with respect to the number of levels assumed and with respect to the strengths of the level-to-level couplings in each scheme.

Of course, these results apply only to the thermodynamic systems described below, but the implications may be the same for many biological systems which have or are assumed to have hierarchical structure. Consider for example a hierarchical predator–prey ecosystem in which levels are defined by species and autonomy of level structure means that the behaviour of a given species can be predicted in terms of variables relating only to that species (with lower and higher levels entering in only as statistically averaged parameters or as constraints). Leaving aside the detailed specification of what constitutes level coupling (level-to-level interactions), Result 1 states that autonomous behaviour is enhanced if smaller species (lower levels) interact more strongly than larger species (higher levels). Result 2 is a caution that a given ecosystem can be modelled in structurally different ways, so that limited input information can have ambiguous implications. On the other hand, as we shall see in examples below, by combining Results 1 and 2 one could in principle discriminate, on the basis of limited input information, between two competing structural schemes.

Models for Result 1

Two hierarchical models were investigated leading to Result 1. In each, behaviour (function) is quantified as thermodynamic behaviour in terms of the free energy F. In the models, one can write down expressions for the exact total free energy in terms of structural parameters; but one can also write down approximate or inferred free energies arising from the assumption, possibly unwarranted, that they depend only on those structural parameters referring to a particular level. To the extent that these inferred free energies decouple in the exact expressions, to this same extent one can say that the level variables behave autonomously.

The first model, illustrated by the example in Fig. 1, is essentially a modified Ising model in which the subunits interact to form dimers at the first level of interaction and the dimers then interact with each other at the second level. Subunits at the lowest level can exist in 2 states (spins up and down in the magnetic analogue) denoted by the variable $\sigma_r = \pm 1$, $r = 1, 2, \ldots, N$, where N is the total number of subunits in the system. States of the system as a whole are 2^N in number, corresponding to the 2 choices of σ_r for each subunit. These 2^N states have energies E_α where $\alpha = 1, 2, \ldots 2^N$ is the state index. The E_α are the eigenvalues of the system Hamiltonian

$$(1)\quad H_N = -\sum_{i=1}^{N/2} [C_1\sigma_{2i}\sigma_{2i-1} + C_2(\sigma_{2i} + \sigma_{2i-1})(\sigma_{2i+1} + \sigma_{2i+2})] - \gamma\Lambda \sum_{i=1}^{N} \sigma_i$$

where C_1 and C_2 are the interaction strengths for subunits and for dimers, and environmental influences are modelled by the last term; Λ is the average external field and γ the coupling strength to the system in question.

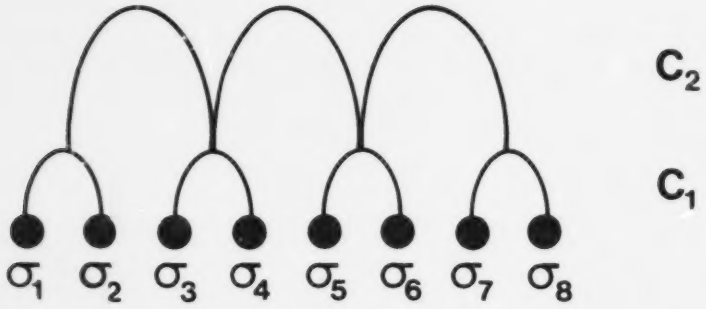


FIG. 1. A structural hierarchy with two levels of interactions. Pairs of subunits interact through field C_1 and pairs of these pairs interact through field C_2. Here $\sigma_{N+1} \equiv \sigma_1, \sigma_{N+2} \equiv \sigma_2$ and $S_N^{(2)}/2 + 1 = S_1^{(2)}$, where N is the number of subunits.

H_N can be rewritten in terms of level variables as

$$(2)\quad H_N = -\sum_{i=1}^{N/2} [C_1 S_{2i}^{(1)} S_{2i-1}^{(1)} + C_2 S_i^{(2)} S_{i+1}^{(2)}] - \gamma\Lambda \sum_{i=1}^{N/2} S_i^{(1)}$$

where the level variables are defined by

$$(3)\quad S_r^{(1)} \equiv \sigma_r, \quad r = 1, 2 \ldots N \quad \text{and} \quad S_r^{(2)} \equiv \sigma_{2r} + \sigma_{2r-1}, \quad r = 1, 2, \ldots N/2.$$

Finally, the system is assumed to be in thermodynamic equilibrium with a heat bath at temperature $T = 1/k\beta$ so that the free energy F_N can be obtained as

$$(4)\quad F_N = -kT \ln Z_N$$

from the canonical partition function

$$(5)\quad Z_N = \sum_{\alpha=1}^{2^N} e^{-\beta E_\alpha}$$

Since all thermodynamic quantities can be obtained from F_N by differentiations, the thermodynamic behaviour is determined once F^N is known. The limit $N \to \infty$ corresponds to the thermodynamic limit in which the system could be said to determine its own temperature.

The second model investigated in TLT (1980) is a hypothetical physical system based on a mathematical construction of a hierarchical model by Dyson (1969) in connection with theorems on ferromagnetic systems. The model with $N = 8$ subunits at level 1 is illustrated in Fig. 2. Essentially subunits at the lowest level form dimers, which pairwise form super-dimers, etc. up to the highest level. As shown in TLT (1980), the Hamiltonian of this system can be written as:

$$H_N = \sum_{p=1}^{n} \left\{ C_p \sum_{r=1}^{2^{n-p}} [S_r^{(p+1)}]^2 \right\} - \gamma\Lambda S_i^{(n+1)}$$

where p denotes the level, C_p the interaction strength at level p, and n the total number of levels defined by $N = 2^n$. At level p, the level variables are defined by $S_r^{(p)} = S_{2r-1}^{(p-1)} + S_{2r}^{(p-1)}$, where r numbers the blocks or subunits at that level. This Hamiltonian can be conveniently rewritten as

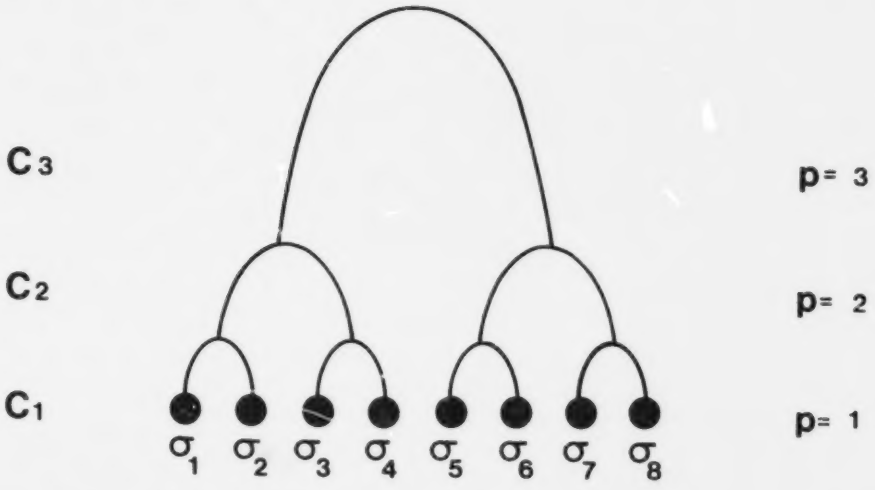


FIG. 2. Dyson hierarchical model with eight subunits and three levels of interaction.

$$H_N = \sum_{ij}^{N} T_{ij}\sigma_i\sigma_j - \gamma\Lambda S_i^{(n+1)}$$

where, for the Dyson model, T_{ij} is a matrix with elements

$$T_{ij} = \sum_{k=P_{ij}}^{n} C_k$$

where P_{ij} is the level index for the lowest level at which the elementary units σ_i and σ_j first interact.

Details for both models are given in TLT (1980) along with the analysis leading to Result 1 above. In effect, functional or behavioural independence at a given level of these hierarchical systems requires the coupling or interaction giving rise to units at that level to be much stronger than the coupling of these units into higher level structures, i.e. C_m/C_{m+1}, where m is the level number, is a strongly decreasing function of m. While the systems considered here are simple physical systems, we anticipate that similar conclusions would follow from an analysis of structure–function relationships in biological hierarchies such as ecosystems.

Model for Result 2

In TT (1982) a modification of Dyson's heirarchical model is studied in some detail. The modification, called the symmetrized Dyson model (SD), is illustrated for the case of $N = 8$ subunits in Fig. 3. Basic subunits are located in a ring and interact with one of their neighbours to form 2-subunit dimers. These interactions are shown as solid lines in Fig. 3(a). The dimers then interact via the secondary interactions depicted as dashed lines in Fig. 3(b). Finally, the 4 subunit tetramers interact via tertiary interactions, depicted as the wiggly lines in Fig. 3(b). The 4 independent configurations possible for the $N = 8$ model are shown in Fig. 3(b). The interaction strengths between these oligomeric units are denoted as C_1, C_2, and C_3.

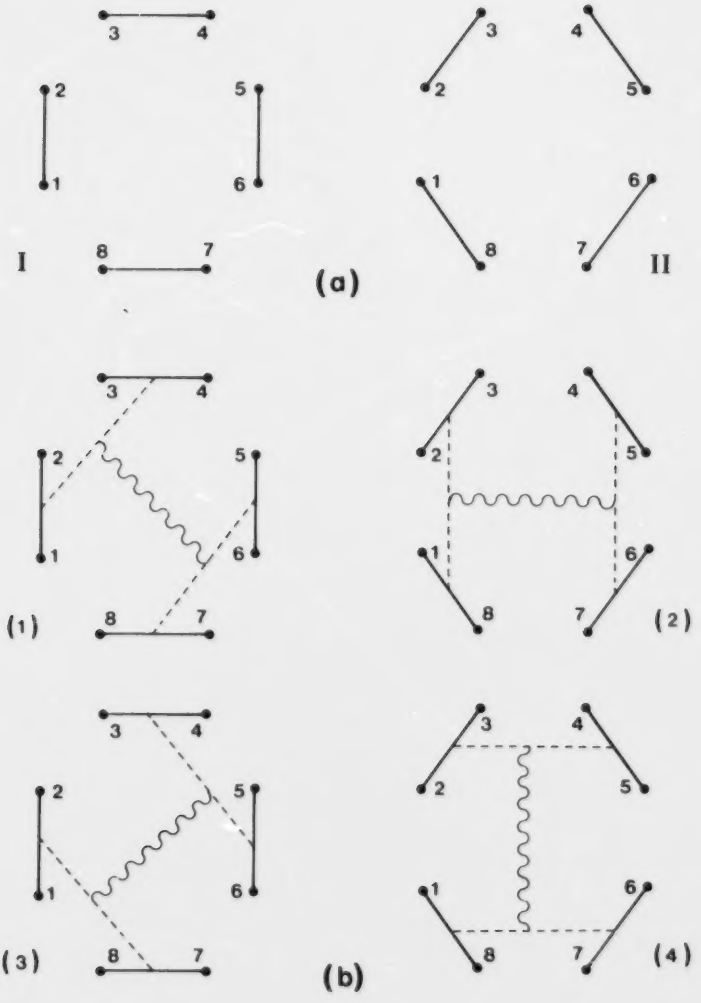


FIG. 3. Bond organizations of Dyson type for a ring of eight subunits. (See TT (82) for details). (a) Bonding patterns I, II arising from interactions at level 1; (b) Bonding patterns 1, 2, 3 and 4 arising from adding in interactions at levels 2 and 3 (dashed and wavy lines, respectively).

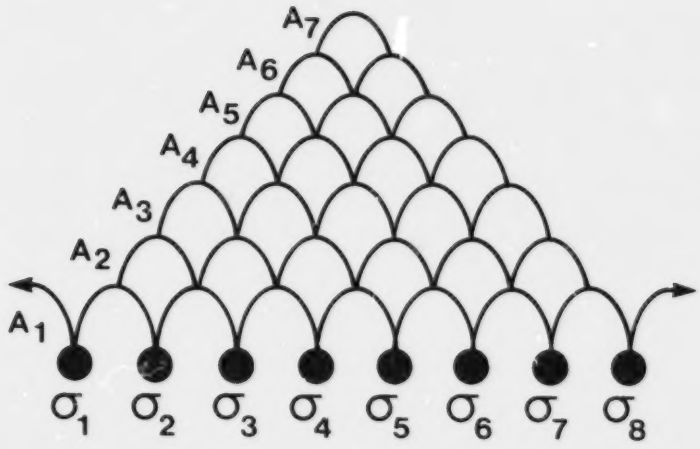


FIG. 4. Nearest neighbour cluster (NNC) model (see TT (82)) with eight subunits and seven levels of interaction. At every level, interaction lines at each end joint together to form a closed loop of subunit–subunit couplings.

It is then shown in TT (1982) that exactly the same thermodynamic behaviour as occurs in this system can be achieved in a variety of other structural organizations of the type shown in Fig. 4, such as the 5 shown in Fig. 5, provided the interaction strengths are appropriately chosen — even the number of levels may be varied. In other words, even when we known the number and character of the elementary subunits in the system and can, in principle, measure thermodynamic behaviour with as much accuracy as desired for the system as a whole, we cannot infer the substructure of interactions and levels. The reader is referred to TT (82) for details.

The implications for ecosystems is that inference from physical and biological measurements at some level of description is very much model dependent and that caution should be used in drawing strong conclusions from what amounts to limited input information.[1] On the other hand, limited input information can be used to discriminate between models in some cases. Result 1 requires graded interactions between successive levels for functional decoupling. Applied to the presumed structures in Fig. 5, structure (3) is consistent with graded coupling but structure (4) is not. Of course, we have used thermodynamic analogues which have no direct relevance to any ecosystem; the conclusions drawn from these examples, nonetheless, are likely to apply with even greater force to more complicated systems with more complicated structure–function relationships than the simple Ising-like models used in this work.

Conclusions

The possible applications of statistical mechanics to ecosystem analysis is a subject in its infancy. The theory may already be useful in isolated instances, but much more work

[1]In this regard an essay of Poincaré written in 1900 and entitled, "Relations between Experimental Physics and Mathematical Physics", may be of interest (reprinted in Nye (1984)).

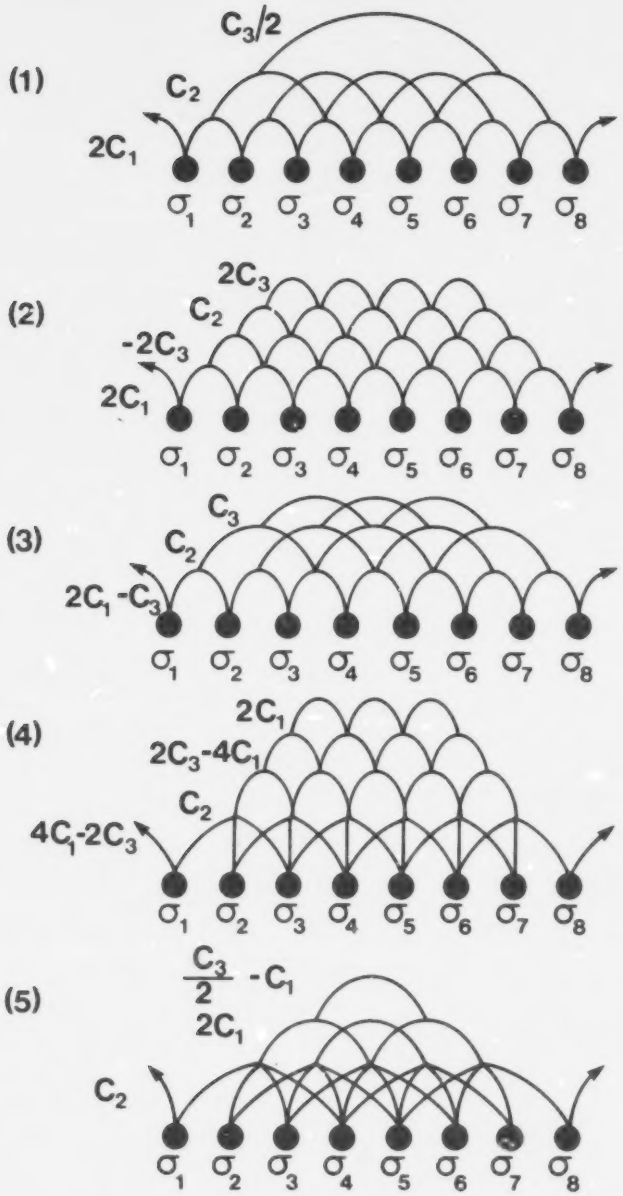


FIG. 5. Five bond structures with NNC-type Hamiltonians which are mathematically equivalent to the Hamiltonian of the symmetrized Dyson (SD) model. Interaction lines at every level join one another at each end as in Fig. 4.

is required to test whether statistical mechanics in some new or adapted form can be usefully employed in analyzing data and constructing models. Nonetheless, statistical mechanics, as successfully used in simple physical examples, provides a "role model" for more complicated situations and a backdrop for inspiration and example.

Acknowledgments

Much of the background work for this paper was research done with financial assistance from the Natural Sciences and Engineering Research Council of Canada. Particular thanks are due to Mr. John Totafurno for useful discussions and to Ana Sousa for word processing.

References

DIRAC, P. A. M. 1950. Generalized Hamiltonian dynamics, Can. J. Math. 2: 129.

GLANSDORFF, P., AND I. PRIGOGINE. 1979. Thermodynamics of structure, stability and fluctuations. Wiley Interscience, New York, NY.

GOODWIN, B. C. 1976. Analytical physiology of cells and developing organisms. Chapter 7. Academic Press, London.

KERNER, E. H. 1972. Gibbs ensemble: biological ensemble. Gordon and Breach, New York, and references therein.

LANDAU, L.D., AND E. M. LIFSCHITZ. 1958. Statistical physics. Pergamon Press, London.

LUMSDEN, C. J., AND L. E. H. TRAINOR. 1979a. On the statistical mechanics of constrained biophysical systems. J. Stat. Phys. 20: 657.

1979b. Non-equilibrium ensembles of self-organizing systems. Can. J. Phys. 57: 23.

1980. On the Hamiltonian content of biodynamics. Bull. Math. Biol. 42: 327.

1984. Hamiltonian statistical mechanics and biological order: problems and progress. Mathematical essays on growth and the emergence of form. University of Alberta Press, Edmonton, Alta.

NYE, M. J. 1984. The history of modern physics 1800–1950. Tomash Publishers, Los Angeles/San Francisco, CA.

PATTEE, H. 1973. Hierarchy theory: the challenge of complex systems. George Braziller, New York, NY.

THOMAS, L. 1975. The lives of a cell: notes of a biology watcher. Bantam, New York, NY.

TOTAFURNO, J., C. J. LUMSDEN, AND L. E. H. TRAINOR. 1980. Structure and function in biological hierarchies: an Ising approach. J. Theor. Biol. 85: 171.

TOTAFURNO, J., AND L. E. H. TRAINOR. 1982. Structure and function in biological hierarchies: ambiguity in the relationship. Preprint, University of Toronto, Toronto, Ont.

TRUBATCH, S. L., AND A. FRANCO. 1974. Canonical procedures for population dynamics. J. Chem. Phys. 33: 1338.

TURING, A. M. 1952. The chemical basis of morphogenesis. Phil. Trans. R. Soc. London B237: 37.

III. ATAXONOMIC AGGREGATIONS

Structure of the Marine Ecosystem: Its Allometric Basis

TREVOR PLATT

Department of Fisheries and Oceans, Marine Ecology Laboratory, Bedford Institute of Oceanography, Dartmouth, Nova Scotia B2Y 4A2

1. Historical Preamble

Appreciation of the value of thinking about the ecosystem in terms of the size-spectrum of the organisms that compose it can be traced to Elton (1927). He devoted an entire chapter of his textbook on ecology to "the animal community" in which the conceptual thread was, explicitly, organism size. Elton's foresight was remarkable: in this one chapter he formulated all of the principles on which the modern theory of the pelagic ecosystem is based. He noted that animal communities can be conceived of as a series of interconnecting food chains forming a food web; he noted that smaller animals are much more abundant than larger ones, and characterised this as the "pyramid of numbers"; he noted that the size of an organism's food bore some relationship to its own body size, and concluded that the elemental food chains comprising the food web are arranged in order of increasing organism size, with the result that the general flow of material through the community is from small organisms to large ones; he noted that, because of their small size, small organisms reproduce faster than large ones; and finally he observed, in a cryptic but pregnant final sentence, that the biomass of prey necessary to sustain a given predator was a function of both the food requirements of the predator and the generation time of the prey.

The other remarkable feature of Elton's chapter was that he introduced these ideas in a deliberate attempt to simplify the bewildering complexity that ecological communities present to the researcher. As such it made the first step towards modern ecosystem analysis. An impression of the extent of the complexity of the pelagic ecosystem itself had been given a few years earlier by Hardy (1924). He summarised the feeding relationships of the herring (*Clupea harengus*) at various stages of its life-history, and showed that different pathways through the food-web would be emphasised at different times, according to seasonal and local fluctuations in the species composition of the plankton. Moreover, the relationship of the herring to its food web was seen to change as it passed through the sequence of ontogenetic stages.

Against this background, the introduction of the trophic level formalism by Lindeman (1942) and by Hutchinson (MS cited in Lindeman, 1942, p. 402–403) might be considered in some respects to have been a retrograde step (Cousins 1980). Although notable as an early (the first?) attempt at *quantitative* analysis of trophic relationships in ecosystems, their approach demanded that organisms be assigned to one, and only one, of a very small number of mutually-exclusive trophic categories, a feature that was to impose severe limitations on the range of conclusions that could be drawn (cf. Rigler 1975). However, for the present context, the paper is notable in that it discusses the transformation of the Eltonian pyramid (of numbers) into a biomass pyramid, and in that it states explicitly that the trophic structure of an ecosystem is controlled in some way by the magnitudes of various physiological rates of the component organisms. Furthermore, the differential equation from Hutchinson's unpublished manuscript is exactly analagous, if the trophic level index is replaced by organism size as a continuous variable, to that used by Platt and Denman (1977, 1978) in developing their theory of the pelagic ecosystem, and later generalised by Silvert and Platt (1980) as the von Foerster, master equation of trophic dynamics.

Returning to the period in which Elton and Hardy were publishing their important work, Haldane (1928) had recognised that the size of an organism was perhaps its single most significant attribute, containing much latent information about its physiology. Considerable time elapsed, however, before these insights were quantified and put into a format that was of direct use to ecologists. Kleiber (1947) and especially Hemmingsen (1960) compiled available data on the dependence of rate of standard respiratory metabolism on organism size (from the smallest unicells to the largest homoiotherms) and found evidence of great regularity.

Quantification of Elton's observation that generation time increases with organism size was to come even later. Bonner (1965, fig. 1) appears to have been the first to make this explicitly clear. Again, he found evidence for remarkable regularity; but he was content to display his findings in pictorial form, stopping short of producing an equation that would summarise them.

Fenchel (1974) produced the first paper in which the size-dependence of both respiration rate and of growth rate were treated as a common theme. With these relationships quantified in operational form, all the necessary pieces were available for development of continuous theories of organism size distribution in the pelagic ecosystem. But it is a sobering thought that 50 years were to have elapsed since the publication of Elton's textbook before such a theory came into being (Platt and Denman 1977, 1978). The essential elements were all there in the 1920's. If further stimulus was required, it surely was offered by the Hutchinson-Lindeman formalism published in 1942. The further (scarcely credible) time delay of 30 years is probably to be explained, paradoxically, by the stifling effect that the trophic level concept had on the development of ecological thought (Cousins 1980).

2. Some Remarks on the Size-dependence of Physiological Rates

The specific rate R' of any metabolic process R is related to organism size, w, through an allometric equation (see review by Calder 1981):

$$(2.1) \quad R' = \frac{R}{w} = \alpha w^{\gamma - 1}$$

The label "specific" implies a normalisation such that R' carries the dimensions $[T^{-1}]$. Body size has dimension $[M]$. The coefficient α, dimensions $[T^{-1}]$, is seen to be just a scale factor. The shape of (2.1) is controlled by the exponent γ (dimensionless). It is found that $\gamma < 1$.

Considerable research has been invested in determination, and rationalisation, of the magnitude of γ for various physiological processes. Comparatively little attention has been given to the magnitude, or interpretation, of the coefficient α. Fenchel (1974) showed that, for specific respiration, unicellular, heterothermic and homoiothermic animals each had characteristic values of α in the approximate ratios 1:8:225. But within the precision of the available data, they shared a common value of the exponent γ.

Corresponding differences between organism types, but of dissimilar relative magnitudes, were found for the allometric dependence of reproduction rate. In this case, the coefficients α for unicells, heterotherms and homoiotherms stood in the approximate ratios 1:2:4. Again, all groups shared a common value of γ. But the most remarkable thing about Fenchel's (1974) analysis was that, although the respiration and reproduction data were compiled from independent sets of studies, the magnitude of γ was the same for respiration as it was for reproduction. This result was so suggestive that Platt and Silvert (1981) concluded that it arose from the fundamental dimensionality of physiological processes, and that it gave evidence for the existence of a universal, size-related time-scale

in organisms to which all processes that could be expressed as a quantity with dimensions $[T^{-1}]$ were related in a simple way.

From the point of view of what follows, the precise magnitude of γ will be of much less significance than is the fact that diverse physiological rates follow similar allometric laws for which the exponents can be identified with some fundamental scale.

3. Data on Size Structure of Pelagic Ecosystem

Relatively few data are available on the distribution by size of organisms in the pelagic ecosystem. None are complete, since no single technique exists to enumerate organisms by size over the entire size range of all pelagic organisms. The most detailed spectra cover the range occupied by the smaller plankton. This range can be sampled either by light microscope or by automatic, inductive counting devices. Neither method is perfect. The best size spectra available are those of Beers et al. (1975, 1982), made with the light microscope. They cover the range roughly from 2 to 250 μm. It is at the small end of this range that the spectra become unreliable. First, because it is now known that there exists in the pelagic zone a large and active biomass of organisms smaller than 2 μm (Johnson and Sieburth 1979, 1982; Waterbury et al. 1979; Li et al. 1983; Platt et al. 1983); and second, because even for organisms of size 2 μm or larger, an unknown proportion of the more fragile types do not withstand the fixation procedures used in processing the samples for analysis.

Inductive counting devices (eg. the Coulter counter) cover roughly the same size range. They too are less reliable at the smaller end of the range, this time because the counts there are difficult to discriminate from instrument noise. Particle spectra of this type are also contaminated across all size classes by an unknown proportion of particulate material other than intact organisms: detrital material that would be ignored in microscope counting.

The earliest particle spectra for a variety of marine regimes were published by Sheldon et al. (1972, 1973). In the particular way that these authors plotted the data, the distributions were more-or-less flat, that is they were lacking in structure. This prompted Sheldon et al. (1972) to speculate that the same flatness, or lack of structure, might persist out to the limits of the pelagic size spectrum, for which detailed data were unavailable.

In fact, this speculation offered nothing that was not available in Elton (1927): if there is a pyramid of numbers, and if it is smoothly sloping, it is always possible to find a simple logarithmic transformation that will map the pyramid into a rectangle. The authors insisted on global flatness in the pelagic size spectrum even against their own evidence (Sheldon et al. 1972, fig. 12). These criticisms aside, it is fair to say that the publication of these spectra (Sheldon et al. 1972, 1973), with the implied rediscovery of the Eltonian pyramid, did serve to rekindle interest in the development of a trophic dynamic theory of the pelagic ecosystem.

It is a matter of record that the idea of overall flatness in the aquatic food chain had been anticipated by H. T. Odum (1971, fig. 3.6d). Furthermore, he made an analogue simulation model based on the size-dependence of metabolism to derive explicit values for the size distribution of the biomass. It was his stereotypic example for food chains arranged in order of increasing organism size. For these reasons it would be fitting to call the biomass spectrum (that is $b(w)$, defined in eqn. 4.1) in ecosystems the Odum–Elton spectrum.

4. Theories of the Pelagic Ecosystem

Dickie (1972) appears to have been the first to attempt to recover the size-dependence of a population property by other than strictly empirical methods. Building on his earlier

work (Paloheimo and Dickie 1965, 1966) he used the illustration in Bonner (1965) to derive the dependence of population generation time on organism size, and recognised that the inverse of this equation was an estimator for the production to biomass ratio in the steady state. Although Dickie's main analysis was in terms of Lindeman trophic levels, he realised (p. 216) that size-dependent relationships might hold the key to an escape from such an unsatisfactory theoretical structure into one based on the continuous (biomass) variable. These were major advances in thinking.

An early attempt to rationalise the results of Sheldon et al. (1972, 1973) in terms of the size-dependence of physiological rates was made by Kerr (1974). This theory however, like the rationalisations of Sheldon et al. (1972), was based on the obsolete (Rigler 1975) concept of trophic levels and would therefore be of only transient interest (Cousins 1980).

Kerr's theory also required that there be a fixed size ratio between a predator and its prey, say a factor of 10 in body length or 10^3 in weight (Kerr 1974). Given that, in the sea, the autotrophs themselves cover a dynamic range in body weight of at least 10^7, it is easily seen that the trophic level assumption represents an unjustified simplification for the pelagic ecosystem.

Progress towards development of a continuous model depended on leaving behind the notion of discrete trophic levels, and on breaking away from the particular way in which Sheldon et al. (1972) presented their data.

With this in mind, Platt and Denman (1977, 1978) introduced the normalised biomass density function $\beta(w)$ such that

$$(4.1) \quad b(w) \equiv \beta(w)\, dw$$

is the mass of organisms, per unit volume of water, with nominal weights in the interval $\{w, (w + dw)\}$.

The approach used was to analyse the flow of biomass energy through the pelagic system as it migrated from particles of small size to those of larger size, the turnover of material within each size class being controlled by the rates of reproduction and respiration characteristic of organisms of the nominal weight characteristic of the size class.

This gave an expression for $\beta(w)$ in the form

$$(4.2) \quad \beta(w) \approx w^m$$

where the exponent m is determined by the parameters of the allometric equations for respiration and growth. Platt and Denman (1978) assigned the value $m = -1.22$. The biomass size spectrum then had the form

$$(4.3) \quad b(w) \approx w^{m+1} \approx w^{-0.22}.$$

This equation was shown (Platt et al. 1984) to be a good fit to the data of Beers et al. (1982), collected in the central gyre of the north Pacific, which indicate an exponent of -0.23 ± 0.03 for the 0–20 m layer and 0.20 ± 0.02 for the 100–120 m layer.

Tseytlin (1981a, b, c) has followed a somewhat similar approach to the analysis of size-distribution data for pelagic organisms.

Cousins (1980) has discussed the general problem of modelling the trophic continuum. He criticises the Lindeman formalism for being a description of the history of energy flow into a particular state: Cousins contends that trophic transfers are rather Markovian, that is the probabilities for trophic transfers are a function of the present state itself, and not of the route by which the energy arrived in that state. Cousins also emphasises that all the basic attributes of a good theory of the trophic continuum were included in Elton's (1927) description of the structure of the ecosystem.

5. Difficulties in Practical Application

The essential difficulty in applying allometric theories of the size structure of the pelagic ecosystem is in the selection of parameters. It is found that the scale coefficient α in eqn. (2.1) is more critical than the exponent γ for deductions about ecosystem properties. However, α is more subject to errors, particularly systematic errors. One problem is that α is temperature-dependent. In general it can be adjusted for differences in temperature according to standard physiological methods, if its magnitude is known accurately at any one temperature. The major snag is that knowledge about α is derived almost entirely from laboratory studies where the organisms are usually maintained under optimum conditions. The application of these data to field populations is contentious. For example, one school of thought maintains that bacteria in the sea are metabolically inert most of the time, perhaps because of low concentration of substrates favourable for growth. In this case, their respiratory expenditure would be considerably below that predicted from eqn. (2.1) using a value for α measured at maximal growth rate. Similarly, Banse (1982a) has concluded that the realised specific growth rates of marine ciliates are much lower than their intrinsic potentials, as a result of their ingestion rates' being very low. It will be shown in the next section that the contribution of the smallest organisms is expected to dominate community metabolism. It is therefore particularly important that systematic errors of this kind in α should be kept as small as possible. Because the question of the relative activity of marine microorganisms has not been resolved, the absolute value of α appropriate to pelagic models in this part of the size spectrum must remain an open question. The dynamic range of α for respiration is greater than that for growth, but respiration is found to control the slope of the pelagic size spectrum (Platt and Denman 1978).

Within the context of the size distribution of biomass in the sea, Azam et al. (1983) have analysed the ecological role of heterotrophic microorganisms. They conclude that about 25% of the net primary production passes via dissolved organic matter through bacteria, which are then grazed by microzooplankton (principally heterotrophic flagellates in the size range 3 to 10 μm). This pathway, called the "microbial loop" can be viewed as an example of material flowing in the trophic continuum against the general direction of small to large. Microzooplankton, through their excreted metabolites, thus become implicated as playing a major part in the remineralisation of nutrients in the sea, a role traditionally assigned to bacteria. It is therefore clear that carbon flow is tightly coupled to the flow of other essential elements.

In formulating energy budgets for individual metazoan organisms, it is usually assumed, as epitomised in the classical Winberg equation, that assimilated ration is channelled into either respiration (maintenance + activity) or growth, and that these processes are otherwise physiologically independent. It is a view that has been challenged by Parry (1983) who holds the opinion that a significant proportion of an organism's heat production arises from the work of biosynthesis. He concludes that many external factors (such as temperature) known to influence respiration exert their effect through the growth rate. That is, the first order effect of perturbation is on the growth rate, to which metabolic rate is scaled. The effect on metabolism is then indirect and second order, and not direct as conventionally believed.

These considerations will have no effect on the estimated slope of the biomass spectrum provided that the size-dependent respiration equation used in its derivation is based on total heat output by growing organisms. They are of interest, however, in that they afford a possible connection with the work of Zotin and Zotina (1967) who analysed organism growth in terms of the theory of irreversible thermodynamics. The indirect effect of

various factors on respiration through growth would then correspond to the existence of an Onsager phenomenological coefficient connecting growth and metabolism. A corollary would be the existence of an antisymmetric coefficient connecting metabolism and growth.

A further difficulty in applying the biomass spectrum relates to a point already discussed (Section 2). That is, the coefficient α increases strongly with the level of organism complexity. If different parts of the size spectrum were characterised by organisms of only one grade of organisation, abrupt changes in spectral slope would occur at the boundaries as a new value of α took over. In fact, the situation is more complex than this, because organisms of different structural complexity coexist in various parts of the spectrum. Banse (1982b) concludes that extremely small metazoa may have a value for α intermediate between that for protozoans and that for poikilotherms, due to their small size per se rather than to their evolutionary complexity.

A final difficulty relates to inadequate knowledge of abundance and biomass in the small size end of the spectrum. Both microscope and Coulter counter spectra cut off around 2 μm. Recent evidence (Li et al. 1983; Platt et al. 1983) indicates that in the subtropical oceans about 50% of the active, autotrophic biomass may be in the form of cells smaller than this. The microheterotrophic biomass, which also extends to sizes smaller than 2 μm, is also imperfectly censused, particularly on a routine basis.

The aggregate of all these difficulties is compounded by the fact that they all bear most heavily on the small-size end of the spectrum. This, as will be shown next, is exactly the part of the spectrum for which we should like to have the strongest data.

6. Size Dependence of Ecosystem Properties

It is worth emphasising that the theory of ecosystem structure embodied in eqn. (4.2) is based on the allometric relationships that apply to individuals and to individual populations. We can use it to derive results about metabolic properties at the community level. For example, community respiration, C_R, is simply the product of the (size-dependent) respiration multiplied by the density function and integrated over the entire spectrum. This is

$$(6.1) \quad C_R = \int_{w_o}^{w_f} R'(w)\beta(w)\,dw$$

where w_o and w_f are, respectively, the smallest and largest size classes represented in the system. This leads to the result (Platt et al. 1984)

$$(6.2) \quad C_R = \frac{\alpha\beta_o}{m+\gamma} w_o^{\gamma}[(w_f^*)^{\gamma+m} - 1]$$

where w_f^* is a dimensionless weight, (w_f / w_o), scaled to the lowest size class w_o. The quantity C_R is to be interpreted as the total respiration of all organisms larger than w_o and smaller than w_f. The terms inside the square bracket are dimensionless, so that C_R has the required dimensions of a rate of respiration. Furthermore, for $w_f^* \gg 1$, we find

$$(6.3) \quad C_R \approx 2\alpha\beta_o w_o^{\gamma}$$

because $(\gamma + m) \approx -0.5$. This approximation is good to 1% when $w_f^* \approx 10^4$.

The cumulative community respiration is therefore dominated by the smallest size classes. This will be true for any spectral shape that is a monotonically decreasing function of w. Note that eqn. (6.2) is scaled to both β_o and γ, illustrating the importance of knowing both the abundance and relative metabolic activity (relative to that under optimal growth

conditions) of the smallest size classes in the plankton. It is shown in Platt et al. (1984) that inclusion of the respiratory contribution of organisms smaller than 2 μm could lead to an increase in the estimate of community metabolism by a factor of three. This range of uncertainty in C_R is a direct reflection of uncertainty in the values of β_n and α for the smallest unicells.

Various authors have collated data from diverse field studies and analysed them statistically in the hope of discovering generalisations about the energy budgets of animal populations (Englemann 1966; McNeill and Lawton 1971; Humphreys 1979; Banse 1979; Banse and Mosher 1980). It has been shown by Damuth (1981), Platt and Silvert (1981) and Lavigne (1982) that these results, won at the cost of great labour, could all be recovered by elementary dimensional analysis using eqn. (2.1) as a fundamental axiom.

All components of the energy budget, for an individual or population, that enter the Winberg equation as additive terms must have the same dimensions as the left hand side of eqn. (2.1) and must therefore have the same dependence on body size as the right hand side of eqn. (2.1). It also follows that any ratio of such terms, computed to derive an efficiency, should be independent of body size. This is the explanation for the insight of Kleiber (1975) that efficiency of food utilisation appears to be independent of body size, and the empirical results of Banse (1979) that net growth efficiency is independent of body size, and the conclusion of Humphreys (1979) that production efficiency does not depend on body size.

In the same way, it is easy to see that the conventional P/B ratio must scale on weight in the same way as the specific respiration rate. That is

$$(6.4) \quad P/B \approx w^{\gamma - 1}$$

This result has been confirmed by Farlow (1976), by Banse and Mosher (1980) and by Lavigne (1982).

7. Size Structure of Benthic Communities

In general, study of the size-dependent structure of benthic communities has lagged behind that of pelagic ecosystems. An early paper was that of Thiel (1975) on the deep sea benthos. The first contribution to discuss the results in terms of modern ideas about biomass spectra was that of Schwinghamer (1981). In that paper, and especially in the more recent one by Warwick (1984), a far higher level of statistical rigour was achieved in the presentation of the data than in anything that has been published on pelagic size structure. It is perhaps fair to say that, on the basis of these two studies alone, the size-dependent structure of benthic communities is better known and better understood than that of pelagic ones. Because of the careful and thoughtful approach used by these two authors, evidence for consistent and repeatable features in the structure of benthic communities began to emerge very early in its exploration.

It is instructive to treat first the results of Warwick (1984). He studied the distribution by size of numbers of *species* of benthic organisms from various shallow-water locations. He found the resulting patterns to be highly conservative. The salient feature of the species spectra is that they are bimodal, with the species minimum between the two peaks occurring at a very repeatable characteristic size (about 45 μm equivalent spherical diameter).

The peaks correspond to previously-recognised biota, the meiofauna and the macrofauna. Warwick therefore speculates that the two species peaks reflect evolutionary optimisation, with respect to body size, of two, fundamentally different modes of life. At the primary sediment–water interface, where space in the plane is a limited resource, the

larger (macrofaunal) peak represents the optimal body size (≈3.1 mg) for the sessile habit. The meiofaunal peak (0.64 μg) represents optimisation to a motile habit at the second-order surface (that between the individual sediment grains and the interstitial water).

It is particularly interesting that the size-dependent relationship for generation time shows that the location of the minimum in species number between the meiofaunal and macrofaunal peaks (≈45 μg) should correspond to a generation time of about one year. Warwick (1984) lists other features of the life history of shallow water benthic organisms that change abruptly, with increasing organism size, at a characteristic size of about 45 μg. These include larval habit and dispersal (benthic or planktonic); semelparous or iteroparous reproduction; discriminate or indiscriminate feeding; asymptotic or continuous growth; and motility or sessility.

Schwinghamer (1981), studying the size-distribution of biomass in benthic, shallow-water communities, also found a characteristic minimum in the spectra at a size between 500 and 1000 μm; consistent with the later results of Warwick (1984). Also consistent were the repeatable occurrences of meiofaunal and macrofaunal biomass peaks. Schwinghamer (1981) also found a local biomass minimum at about 10 μm, a size that separates unicellular forms (bacteria and algae), living on the surface of sediment grains, from the mobile interstitial fauna.

The aggregate of this recent research on the benthos should serve as a caution, in the analysis of pelagic systems, that subtleties associated with life history phenomena can complicate the work of describing size spectra according to simple allometric relationships (see also Banse 1982b).

8. Future Applications of Biomass Spectra

Three principal arguments suggest that the biomass spectrum has an important future in marine ecology. These are, first, that it provides an operationally viable, alternative taxonomy that is at least as cost-effective as the conventional one; second, that it lends itself to determination by automatic devices including those deployed non-destructively from a moving ship; and third, that it contains latent information about community metabolism that cannot be obtained from a conventional taxonomic description.

Some practical problems with classical taxonomy have been listed by Bahr (1982). Its premises are obsolete, having been established before the theory of evolution; many existing species have yet to be named; at best, its application in nature is subjective; it requires very specialised training to apply; and it is inappropriate for many organisms, such as clones of microorganisms.

Two other weaknesses are of direct relevance here. One relates to a further aspect of its obsolescence: it is out of step with that school of ecological thought that attaches importance to the macroscopic view and recognises that organisms must obey basic thermodynamic laws. The second is simply that a conventional taxonomic description of a typical pelagic sample is highly demanding on time. And not to put too fine a point on it, after the job is done the investigator is not always certain what to do with the result.

A description in terms of size distribution of the biomass goes a long way towards providing a taxonomy of more immediate ecological utility. It is attractive also because many parts of the size spectrum can be assessed automatically by relatively unspecialised personnel. This is true of living samples, preserved samples or even of undistrubed samples in situ. This is not to say that careful work is no longer necessary, as the elegant studies of Schwinghamer (1981) and of Warwick (1984) attest. Modern ways to size particles include electrical (resistive), acoustic, optical techniques and combinations of techniques.

Once a size-spectrum has been constructed for a particular sample or station, we already have considerable potential information about the physiology of the community, and the spectrum suggests its own applications in a way that a conventional taxonomic description does not. Therefore, it is worthwhile to consider basing pelagic ecosystem studies on size-dependent principles. This has yet to be achieved for any given site.

It can be argued that ecosystem description in terms of size only goes too far in the direction of generality. Thus Bahr (1982) and Cousins (1980) would favour some additional classification according to functional trophic groups. For example, as a minimum requirement, Cousins would separate out detritus and autotroph particles from the rest, and Bahr would go much further. But the beauty of a simple biomass separation is that it is realisable in operational terms with more-or-less unequivocal results.

Acknowledgements

I thank S. H. Cousins and L. M. Dickie for constructive criticism of the manuscript.

References

AZAM, F., T. FENCHEL, J. G. FIELD, J. S. GRAY, L. A MEYER-REIL, AND F. THINGSTAD. 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10: 257–263.

BAHR, L. M. 1982. Functional taxonomy: An immodest proposal. Ecol. Mod. 15: 211–233.

BANSE, K. 1979. On weight dependence of net growth efficiency and specific respiration rates among field populations of invertebrates. Oecologia. 38: 111–126.

1982a. Cell volumes, maximal growth rates of unicellular algae and ciliates, and the role of ciliates in the marine pelagial. Limnol. Oceanogr. 27: 1059–1071.

1982b. Mass-scaled rates of respiration and intrinsic growth in very small invertebrates. Mar. Ecol. Prog. Ser. 9: 281–297.

BANSE, K., AND S. MOSHER. 1980. Adult body mass and annual production/biomass of field populations. Ecol. Mongr. 50: 355–379.

BEERS, J. R., F. M. H. REID, AND G. L. STEWART. 1975. Microplankton of the North Pacific Central Gyre. Population structure abundance. June 1973. Int. Revue. Ges. Hydrobiol. 60: 607–638.

1982. Seasonal abundance of the microplankton population in the North Pacific Central gyre. Deep-Sea Res. 29: 227–245.

BONNER, J. T. 1965. Size and cycle. An essay on the structure of biology. Princeton University Press, Princeton, NJ. 219 p.

CALDER, W. A. 1981. Scaling of physiological processes in homeothermic animals. Ann. Rev. Physiol. 43: 301–322.

COUSINS, S. H. 1980. A trophic continuum derived from plant structure, animal size and a detritus cascade. J. Theor. Biol. 82: 607–618.

DAMUTH, J. 1981. Population density and body size in mammals. Nature 290: 699–700.

DICKIE, L. M. 1972. Food chains and fish production. ICNAF Spec. Publ. No. 8, p. 201–221. *In* Environmental conditions in the N.W. Atlantic. International Commission for the North West Atlantic Fisheries Spec. Publ. No. 8.

ELTON, C. 1927. Animal ecology. Macmillan, New York, NY. 207 p.

ENGELMANN, M. D. 1966. Energetics, terrestrial field studies and animal productivity. Adv. Ecol. Res. 3: 73–115.

FARLOW, J. O. 1976. A consideration of the trophic dynamics of a late Cretaceous large-dinosaur community (Oldman Formation). Ecology. 57: 841–857.

FENCHEL, T. 1974. Intrinsic rate of natural increase: The relationship with body size. Oecologia 14: 317–326.

HALDANE, J. B. S. 1928. On being the right size, p. 321–325. Reprinted *In* H. Shapely, S. Raffort, and H. Wright [ed.] A treasury of science. Harper (1946), New York, NY.

HARDY, A. C. 1924. The herring in relation to its animate environment, Part 1. The food and feeding habits of the herring. Fish. Invest. Lond., Ser. II, 7, No. 3.

HEMMINGSEN, A. M. 1960. Energy metabolism as related to body size and respiratory surfaces, and its evolution. Reports of the Steno Memorial Hospital, Copenhagen. 9 (part 2) p. 110.

HUMPHREYS, W. F. 1979. Production and respiration in animal populations. J. Anim. Ecol. 48: 427–453.

JOHNSON, P. W., AND J. MCN. SIEBURTH. 1979. Chroococcoid cyanobacteria at sea: A ubiquitous and diverse phototrophic biomass. Limnol. Oceanogr. 24: 928–935.

1982. In-situ morphology and occurrence of eucaryotic phototrophs of bacterial size in the picoplankton of estuarine and oceanic waters. J. Phycol. 18: 318–327.

KERR, S. R. 1974. Theory of size distribution in ecological communities. J. Fish. Res. Board Can. 31: 1859–1862.

KLEIBER, M. 1947. Body size and metabolic rate. Physiol. Rev. 27: 511–541.

1975. The fire of life. (2nd ed.) Krieger, N.Y.

LAVIGNE, D. M. 1982. Similarity in energy budgets of animal populations. J. Anim. Ecol. 51: 195–206.

LI, W. , D. V. SUBBA RAO, W. G. HARRISON, J. C. SMITH, J. J. CULLEN, B. IRWIN, AND T. PLATT. 1983. Autotrophic picoplankton in the tropical ocean. Science 219: 292–295.

LINDEMAN, R. L. 1942. The trophic-dynamic aspect of ecology. Ecology 23: 399–418.

MCNEILL, S. , AND J. H. LAWTON. 1971. Annual production and respiration in animal populations. Nature (London) 225: 472–474.

ODUM, H. T. 1971. Environment, power and society. Wiley, New York, NY. 331 p.

PALOHEIMO, J. E. AND L. M. DICKIE. 1965. Food and growth of fishes. I. A growth curve derived from experimental data. J. Fish. Res. Board Can. 22: 521–542.

1966. Food and growth of fishes. III. Relations among food, body size, and growth efficiency. J. Fish. Res. Board Can. 23: 1209–1248.

PARRY, G. D. 1983. The influence of the cost of growth on ectotherm metabolism. J. Theor. Biol. 101: 453–477.

PLATT, T., D. V. SUBBA RAO, AND B. IRWIN. 1983. Photosynthesis of picoplankton in the oligotrophic ocean. Nature 300: 702–704.

PLATT, T., AND K. DENMAN. 1977. Organisation in the pelagic ecosystem. Helgol. Wiss. Meeresunters. 30: 575–581.

PLATT, T., AND K. L. DENMAN. 1978. The structure of the pelagic marine ecosystems. Rapp. P.-V. Reun. Cons. Int. Explor. Mer. 173: 60–65.

PLATT, T., M. LEWIS, AND R. GEIDER. 1984. Thermodynamics of the pelagic ecosystem: Elementary closure conditions for biological production in the open ocean, p. 49–84. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, London. 733 p.

PLATT, T., AND W. SILVERT. 1981. Ecology, physiology, allometry and dimensionality. J. Theor. Biol. 93: 855–860.

Rigler, F. H. 1975. The concept of energy flow between trophic levels, p. 15–26. *In* W. H. van Dobben and R. H. Lowe-McConnel [ed.] Unifying concepts of ecology. W. Junk, The Hague.

SCHWINGHAMER, P. 1981. Characteristic size distributions of integral benthic communities. Can. J. Fish. Aquat. Sci. 38: 1255–1263.

SHELDON, R. W., A. PRAKASH, AND W. H. SUTCLIFFE JR. 1972. The size distribution of particles in the ocean. Limnol. and Oceanogr. 17: 327–340.

SHELDON, R. W., W. H. SUTCLIFFE JR., AND A. PRAKASH. 1973. The production of particles in the surface waters of the ocean with particular reference to the Sargasso Sea. Limnol. Oceanogr. 18: 719–733.

SILVERT, W., AND T. PLATT. 1980. Dynamic energy-flow model of the particle size distribution in pelagic ecosystems, p. 754–763. *In* W. Charles Kerfoot [ed.] Evolution and ecology of zooplankton communities. The University Press of New England, New Hampshire.

THIEL, H. 1975. The size structure of the deep-sea benthos. Int. Rev. Gesamten. Hydrobiol. 60: 575–606.

TSEYTLIN, V. B. 1981a. Size distribution of pelagic organisms in tropical ocean regions. Oceanology 21: 86–90. (English translation)

1981b. Energy characteristics and size distribution of pelagic organisms in the tropical regions of the ocean. Oceanology 21: 382–387. (English translation)

1981c. Estimating the vertical flow of detritus from the surface of the tropical ocean. Oceanology 21: 508–511. (English translation)

WATERBURY, J. B., S. W. WATSON, R. R. GUILLARD, AND L. E. BRAND. 1979. Widespread occurrence of a unicellular marine plankton cyanobacterium. Nature 277: 293–294.

WARWICK, R. M. 1984. Species size distributions in marine benthic communities. Oecologia 61: 32–41.

ZOTIN, A. I., AND R. S. ZOTINA. 1967. Thermodynamic aspects of developmental biology. J. Theor. Biol. 17: 57–75.

Size and Metabolism in Natural Systems

WILLIAM A. CALDER III

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Introduction

For most of the elapsed half-century since publication of the metabolic equations of Kleiber and of Brody and Proctor (see Kleiber 1961), allometry has been used only in isolation, as a statistical tangent to studies of single variables such as metabolic or heart rate, organ or home range size. The integrating papers of Adolph (1949) and Stahl (1962) have been neglected for the most part. Only recently has there been a significant and widening appreciation that allometry (alias similarity analysis or biological scaling) has a good potential for unification in biology. Unifying principles can, in turn, provide a hypothetical framework for going beyond apparent limitations to reconstruct the paleontological past and to serve as a conceptual platform on the vastness of the sea (Western 1979, 1980; Platt and Silvert 1981; Damuth 1981a, b, 1982; Lavigne 1982; Calder 1983c, 1984).

Allometry is an empirical model employing scaling and dimensional analysis. Scaling and dimensional analysis have been identified by Platt et al. (1981) as crucial for descriptions of biomass, processes, and fluxes in marine ecosystems. Because it has a longer history of application in terrestrial physiology and ecology, I will begin the review there.

Within a class or common body plan, body size is the single characteristic that exceeds all others in determining what an animal needs, what it can exploit, and how much time it has to meet its needs and replace itself. As I will argue below, the entire life schedule is scaled to body size. As we perceive declines in fisheries and the need for moratoriums on harvests to save endangered species, the allometry of productivity and replacement is quite relevant.

The relationship between size (mass, M in kg) and a physiological, morphological, or life history variable (y) is not often in direct linear proportion, so we write:

(1) $y = aM^b$

following Huxley (1932), Brody (1945), and Kleiber (1961). The scaling exponent b is the "other (allo-) measure" of allometry. It shows the quantitative effect of size within a group of animals (phylogenetic or ecological) and is often displayed as the slope of a log–log graph of the function. In the log–log format, the coefficient a is the y-intercept characterizing that group of animals. The correlation coefficients, squared (r^2) tell what proportion of the variance in the data may be attributed to size alone. The r^2 values usually are between 0.67 and 0.99; the average of about 0.75 says that adaptations to other factors account for only one-fourth of the variance.

One must bear in mind that allometric equations are only correlations. As such they explain nothing, but they do provide the clearest patterns which we seek to explain, hopefully a glass-bottomed bridge from the familiar to that obscured by the diversity in the oceans.

Although taking a different tack, this allometric approach will follow the expedience recommended by Steele (1974). "Because we cannot move freely within the deep sea, it is desireable to test our hypotheses by comparison with relevant studies on land." I not only agree with Steele's assumptions of a central importance for energy flow rates in an

ecosystem and of "the existence of fairly broad hypotheses relating populations to their environments", but I believe that allometry strengthens the acceptance of these assumptions.

An ecosystem's energetics are a super-organismic summation of the metabolism of organisms within the ecosystem. The biology of terrestrial mammals has been subjected more extensively to allometric analysis than the biology of other animals. In addition, the ecologically important "marine mammals, like sea birds and marine turtles, have adapted an essentially terrestrial morphology, physiology, and behavior to the medium of the sea" (Ray 1981). Hence I will emphasize mammalian metabolism as the basis for further analysis. I will also follow Stahl (1962) in referring to a size-independent ($\propto M^{0}$) allometric cancellation or product as a "design constant." Bear in mind that these represent only general or average values subject to considerable variation among species as adaptations (or compensations for adaptations) to their specific niches!

A Metabolic View of Life

The basal or standard metabolic rates ($\dot{E}_{st}$, here converted to watts) of animals are largely accounted for by the mass-dependency first described by Kleiber (see 1961):

(2) $\dot{E}_{st} = 3.27M^{0.76}$

It is easier to describe the consequences of this scaling or of a hypothetical linear-scaling ($M^{1.0}$) than it is to explain why we find $M^{1.4}$ scaling (I will use decimal exponents to signify empirical statistics or fittings, and fractional exponents to suggest general patterns). The metabolic intensity or mass-specific metabolic rate,

(3) $\dot{E}_{st}/M^{1.0} = 3.27M^{-0.24}$

tells that a unit mass of elephant flesh requires oxygen, energy, and heat dissipation at slower rates than an equivalent mass of small mammal tissue.

When the level of activity is elevated from the basal level to sustain an interspecifically comparable level of running locomotion, the coefficient may increase an order of magnitude, but the exponents are not significantly changed from the $M^{1.4}$ or $M^{-1.4}$ above. It is useful to view these relationships not only in terms of energy per unit of time (J/s as above) but in their reciprocals of time per unit of energy to express the metabolic pace of an animal. This metabolic pace or physiological time (t_{met}) duration for consumption of a fraction $X/M^{1.0}$ of its body mass in the form of stored fat, at 39.7 kJ/g is:

(4) $t_{met} \propto [39.7(X/M^{1.0})/3.27M^{0.76}] \propto M^{0.24}$

thus t_{met} increases with size. The elephant is doing the same thing as the mouse, biochemically, but at a slower pace. Furthermore, it appears that the entire lives of animals seem to be scaled in proportion to the metabolic or physiological time. Time between heartbeats, times for a breath cycle or cycles of fast muscle twitch, turnover time for body water via renal processing (Calder and Braun 1983), and the duration of pregnancy, growth, maturation, and maximum recorded lifespan ($t_{ls,max}$) are proportioned by approximately the same scaling (Lindstedt and Calder 1981). For example (Sacher 1959):

(5) $t_{ls,max} = 11.6M^{0.20}$

This means that in a lifetime a unit of tissue uses about the same amount of energy (E_{total}) to support the standard metabolism (combining equations 3 and 5):

(6) $E_{total} \propto M^{-0.24} \times M^{0.20} \propto M^{-0.04} \approx M^{0}$

In nature, animals are not at the basal level for much of the time, in a life that may not attain the record longevity. The calculated average daily metabolism scales as $M^{2/3}$ (King 1974; Garland 1983b) while life expectancy scales more like $M^{1/3}$ (Damuth 1982; Calder 1983b). In nature, the lifetime product would have the same sort of scaling as in equation (6):

(7) $E_{\text{total}} \propto (M^{2/3}/M^{1/0})(M^{1/3}) \propto M^{0}$

Not only is there a size-independent lifetime metabolism, but there is also an average lifetime reproduction constant for mammals. The mean annual birth rate scales as $M^{-1/3}$ (Western 1979), so that for the $M^{1/3}$ mean expected lifespan, a design constant of $M^{-1/3} \times M^{1/3} \propto M^{0}$ lifetime reproductivity appears to be the rule. These two size-independencies challenge the common assumption that small mammals put proportionately more energy into reproduction while large mammals must invest proportionately more into self-maintenance (Calder 1983c).

Damuth (1981a) determined that the population densities (ρ) scale as $M^{-0.75}$, reciprocal to individual metabolic rates. Thus, in one km^2 of habitat, the metabolic rate of the population of herbivores is proportional to $M^{-0.75} \times M^{0.75} \propto M^{0}$. This design constant indicates that the small species and the large each take the same proportion of the primary productivity. The biomass of this population is the product ρM, scaled $M^{-0.75} \times M^{1.0} \propto M^{0.25}$, so standing crop biomass increases with body size. Per kilogram of this standing crop, the metabolic rate cancels the positive allometry.

The production or conversion efficiency of herbivorous mammals is a size-dependent $1\frac{1}{2}$–3% at least in captivity. In other words productivity of new tissue in growth and reproduction scales like metabolism ($M^{3/4}$ on an individual basis or M^{0} for that population) (Banse 1979; Humphreys 1979; May 1979; Platt and Silvert 1981). Hence the ratio of productivity to biomass would be expected to scale as $M^{-1/4}$. Empirical determinations range from $M^{-0.27}$ to $M^{-0.33}$. The greater negative exponent could be derived from the scaling of field metabolism (see above):

(8) $$\text{productivity/biomass} = \frac{\rho \dot{E}}{\rho M} \propto \frac{(M^{3/4})(M^{2/3})}{(M^{3/4})(M^{1.0})} \propto M^{-1/3}$$

Because productivity can be expressed in terms of mass or energy and the biomass can be used directly or converted to its energetic equivalent, one can cancel out all dimensions in the ratio except time:

(9) $$\text{productivity/biomass} = \frac{\rho(\text{energy/time})}{\rho(\text{energy})} \text{ or } \frac{\rho(\text{mass/time})}{\rho(\text{mass})} = t_{cc}^{-1}$$

The reciprocal is the turnover time for the standing crop (t_{cc}), which would scale as $M^{1/4}$ to $M^{1/3}$. This t_{cc} should be the same as life expectancy at birth, and indeed there is good agreement between some of the independent derivations (Calder 1982, 1983c; Damuth 1982; Caughley and Krebs 1983).

The reciprocal of population density (= per capita space, ρ^{-1}) has the dimensions of area/animal and a scaling of $M^{3/4}$, whereas home-ranges of herbivores, although dimensionally the same, scale as $M^{1.02}$. This indicates a trend with increasing body size for the home range to exceed per capita space, resulting in a spatial overlap of $M^{0.27}$ (derived) to $M^{0.34}$ (empirical, Damuth 1981b) which roughly parallels the scaling of biomass and of life expectancy. The small herbivorous mammal's home range will, in the long run, support approximately the same number of animals as the large mammal's larger home range. However, this is a temporal sequence of the brief turnovers of individual mice as contrasted with a longer-term spacial coexistence favoring sociality in the wapiti (Calder 1982).

Sociality is also favored by the difference in scaling between life expectancy ($\propto M^{1.1}$) and maximum physiological lifespan (equation 5), because the larger the mammal, the greater the average proportion of possible longevity attained. This translates into size-dependent population–age distributions, with the survival of proportionately more wise elders to give younger herd members the benefit of their experience. Allometry is thus woven into sociobiology. Note the similarity to what occurs in the sea: "... because among whales the mortality rates are low and many age groups are present, new recruits are a small fraction of the population. In contrast, the mortality rates in fish stocks are fairly high and the population has few age classes, so that recruits are a large fraction of the population." (Beddington and May 1982).

I noted above that the scaling of metabolic or physiological time is widespread and consistent in reproduction and growth. This appears in the scaling of the intrinsic or maximal population growth rate (r_{max}) which scales as $M^{-0.26}$ to $M^{-0.36}$ (Fenchel 1974; Hennemann 1983; Caughley and Krebs 1983; Schmitz and Lavigne 1984), and from which the minimum time for population doubling (t_{2x_0}) can be derived, $\propto M^{0.26}$ to $M^{0.36}$. Since reproduction and turnover are scaled in this fashion, one might expect the periodicity of boom–crash population irruption cycles to be scaled similarly. Ecologists attempted for years to correlate cycles of voles *or* varying hares with environmental periodicities, but it now appears that the periodicity is endogenous and may not point directly to the cause (Calder 1983a; Peterson et al. 1984). The predominance of intrinsic vs. extrinsic regulation of population density appears to be mostly a matter of body size in herbivorous mammals (Caughley and Krebs 1983).

Finally, it is interesting to note that speed of travel and distances travelled daily by walking and running tend to approximate $M^{1.4}$ scalings, so larger mammals tend to go farther and faster and thereby to quality as the migrators amongst the walkers and runners (Taylor et al. 1982; Garland 1983a, b).

As a reliable generalization, the similarity in scalings to body size of terrestrial, eutherian mammals is overwhelming when the interspecific, adult body size range is wide enough to establish an accurate regression slope. While the slopes or scaling exponents (b) are similar in interclass comparisons, there are often distinct differences in coefficients (a). For example, heterothermic animals are spared the expense of thermoregulation, and are characterized by lower *a*-values. Consequently their size-independent mean production efficiencies tend to be an order of magnitude greater than those of mammals and birds. Compared on an equal-size basis, the productivity/biomass ratios of fish and invertebrates are lower than for mammals, so their turnover time would be longer.

The cost of locomotion (energy per km) is progressively lower for flying and swimming, compared to cursorial locomotion, thus making it practical for small insects, fish, birds, and bats to migrate whereas only large mammals such as caribou undergo seasonal migrations of any great distance (Tucker 1970; Schmidt-Nielsen 1972; Taylor et al. 1982). Now I shall narrow the discussion from these generalities to particular groups of animals relevant to marine biology: invertebrates, fish, marine birds and marine mammals.

Invertebrates

Hemmingsen (1961) showed that the $M^{3/4}$ scaling of metabolism is not limited to homeotherms, to vertebrates, or even to multi-cellular animals, but is fairly ubiquitous throughout the animal kingdom. Within classes treated separately, there is a wider variety in exponents *b* (see Bertalanffy 1957; Banse 1982; Peters 1983). Annual production does not depart significantly from a linear proportionality to respiration (Humphreys 1979), so productive efficiency can be considered to be size-independent in the Mollusca, Crustacea, and other non-insect invertebrates. Productivity/biomass ratios scale as $M^{-0.27}$ to $M^{-0.46}$

(Fenchel 1974; mean exponent -0.34 ± 0.11 in equations of Banse and Mosher 1980). As obtains for intraspecific scaling of metabolism in mammals (Feldman and McMahon 1983), intraspecific scaling of productivity/biomass of invertebrates may exhibit variations in exponents from those in interspecific regressions (see Greze 1978). The standing crops would turn over in times proportioned to $M^{1.3}$. Invertebrate population r_{max} scales as $M^{-0.35}$ to $M^{-0.36}$, essentially parallel to productivity/biomass scaling. Therefore, the reciprocals (t_{sc} and time for population doubling, t_{2N_0}) also share a common scaling; the ratio t_{sc}/t_{2N_0} may be a size-independent design constant.

Body size consequences lead not only to quantitative trends, but to qualitative changes. For example, parental care is a life history trait with several correlates, including body size. Amongst co-occurring benthic marine invertebrates, there is a tendency for greater brood care with smaller adult size. This may be a consequence of any of several factors, such as the allometries of gamete production, brood patch size, and longevity (Strathman and Strathman 1982).

Fish

While the data base is limited, standard metabolic rates of fishes scale in roughly the same fashion as for mammals. Bartholomew (1982) lists $M^{0.70}$ scaling for four species of freshwater fish and $M^{0.85}$ for salmon. Growth rates of fish scale as $M^{0.61}$ (Case 1978). Productivity/biomass shows $M^{-0.26}$ (Banse and Mosher 1980). The productivity was found to be proportional to (respiration)$^{0.83 \pm 0.09}$, not statistically distinct from linearity, so production efficiency is size-independent for fish as for mammals, but higher (Humphreys 1979).

There is an unexplained fluctuation in recruitment of yearling fish to North Sea populations. For example, the haddock recruitment shows peaks that averaged about $3\frac{1}{4}$ years from 1920 to 1970 (Steele 1974). Is there a size-dependent periodicity as occurs in mammalian population irruptions, or does the virtually complete harvest of primary productivity in oceans result in exogenous conditions that overwhelm any endogenous periodicity? Steele and Henderson (1984) modelled fluctuations in fish stocks. One of the four coefficients in their model is intrinsic growth rate (r_{max}). While there does not appear to be an allometric equation for r_{max} specifically derived for fish, the size-dependency of r_{max} in other groups suggests that it would be a factor in fish as well. This seems worthy of further investigation.

Marine Birds

The standard metabolism of birds is quantitatively indistinguishable from that of eutherian mammals (Bartholomew 1982), if one excludes the order Passeriformes (that is without oceanic representatives). Schneider and Hunt (1982) used avian metabolic allometry to estimate the food requirements and carbon fluxes of marine birds in the Bering Sea. They also presented data on bird densities for 10 species, but the size range (65 – 1095 g) is too small to show a distinct body size dependency above the other variability.

Avian production efficiencies are also quantitatively similar to those of mammals. Maximum lifespans of birds are significantly longer than for mammals, but the $M^{1.5}$ scaling is the same (equation 5). Avian life expectancy scales as $M^{0.46}$, but this may be an artifact of the erroneous assumption of age-independent mortality (Calder 1983b). Body mass exponents for several life history traits are indistinguishable in birds and mammals, but, again, the coefficients often differ (Western and Ssemakula 1982). In the widely distributed marine order Procellariiformes (albatrosses, shearwaters, and petrels) the time until first breeding scales as $M^{0.22}$. This fits the general pattern of physiological times noted above (Lindstedt and Calder 1981).

Marine Mammals

The Cetacea includes the largest living mammals. Like many large terrestrial mammals, they are highly social. The structural and physiological allometry of whales differs from terrestrial mammals in regard to the buoyant support of sea water and the breath-holding required for prolonged dives. Skeletal mass is almost linearly scaled for whales ($\propto M^{1.02}$; Smith and Pace 1971) but distinctly hyperallometric in terrestrial mammals ($\propto M^{1.09}$; Prange et al. 1979). Alveolar surface area is less in whales than what would be predicted by extrapolation from terrestrial mammals (Weibel 1982).

Because of the enormous technical problems of measuring oxygen consumption ($\dot{V}_{O_2}$) in whales, the values for their resting metabolism are only estimates, but these do give $M^{3/4}$ scaling (Gaskin 1982, p. 95) so the energy requirements of basal maintenance may not depart significantly from the mouse-to-elephant plot, extended. On the other hand, Kanwisher and Sundnes (1965) measured $\dot{V}_{O_2}$ of 305–340 mL/min by young harbor porpoises (*Phocaena phocaena*, 26 and 31 kg; full-sized adults range 50–75 kg), when "close to, if not in," the range of thermoneutrality. This was 2.6 times the prediction from equation (2). Their metabolic turnover times would therefore run only 39% of a typical terrestrial counterpart. Sexual maturity is attained in half the predicted time, so perhaps they are on a faster physiological time-scale, a matter to be considered below. Direct metabolic-measurements from whales are needed to confirm the $M^{3/4}$ scaling.

Mass-specific metabolic rates ($\propto M^{-1/4}$, equation 3) are sometimes misinterpreted as a measure of "metabolic efficiency" (which they are *not*!) through neglect of the scaling of physiological time (Calder, 1984). For example, Ray (1981) stated ". . . whales are much more efficient than seals, requiring only about 4× body weight of food per season, whereas crabeater seals might require 23× body weight."

In fact, 4 times the mass (93869 kg) of a blue whale and 23× the 220 kg of a crabeater seal put the scaling of the seasonal requirement at:

(10) kg food = $109\ M^{0.71}$

This exponent is quite similar to that in Kleiber's equation. Kleiber (1961) also calculated the efficiency for conversion of hay to body mass gains of rabbits or of a steer. Because the rabbits (or seals) live on a comparatively faster metabolic scale than the steer (or whale), they go through food relatively faster but they also produce meat faster. Thus the physiological or metabolic time scale is a concept needed in marine ecology.

Schmitz and Lavigne (1984) re-examined the scaling of r_{max} and found no justification for excluding marine mammals from consideration with terrestrial species as had been done by Hennemann (1983). As yet, there seems to be no reason to assume that the proportionate allocation for maintenance and reproduction of marine mammals is any different from terrestrial mammals; population doubling time scales in parallel to metabolic turnover time.

Consideration of physiological time has a significant impact on the interpretation of development and maturation in marine mammals. In Ray's (1981) tabulation of marine mammals, the sea otter and four pinnipeds were characterized as having early maturation and moderate to high rates of reproduction. However, when compared with allometric predictions for terrestrial mammals of the same sizes, the gestation periods and maturation times of all but the crabeater seal were actually longer than these predictions. Litter sizes are mostly 1, and gestation periods were too long to produce more than 1 pup per year. On the other hand, the walrus and the four cetaceans listed were characterized as maturing late, but the available information (Bryden 1972; Walker 1975; Slijper 1979; Gaskin 1982) indicates that maturation (and/or gestation period as an index of physiological time) is either on the typical mammalian schedule from allometric predictions (walrus) or considerably ahead of schedule (see Table 1).

TABLE 1. Gestation periods and maturation times as indices of physiological time scales of selected marine mammals. Predicted times are based on equations (listed as footnotes) for terrestrial eutherian mammals. Corrections applied to predicted minimum population doubling times (at r_{max} for terrestrial mammals) assuming that the pendulum theory of life histories is correct (see Calder 1983c, 1984).

Species	Body mass (kg)	Gestation (days)			Maturation (years)				Correction to population doubling (years)		
		Observed	Predicted[a]	Obs/Pred	Observed	Predicted[b]	Obs/Pred	Ray (1981)	(mean Obs/Pred)	Predicted[c]	Corrected
Sea Otter (*Enhydra lutris*)	27	240	156	1.54	3	2.0	1.50	Fairly early	1.52	3.0	4.6
Northern fur seal (*Callorhinus ursinus*)	55	273	188	1.45	3	2.4	1.25	Fairly early	1.35	3.9	5.3
Harbor seal (*Phoca vitulina*)	65	240	196	1.22	—	—	—	Early	1.22	4.1	5.0
California sea lion ♀ (*Zalophus californianus*)	90	342	213	1.60	3	2.8	1.07	Fairly early	1.34	4.6	6.2
Crabeater seal (*Lobodon carcinophagus*)	220	255	255	1.00	2	3.6	0.56	Early	0.78	6.4	5.0
Walrus (*Odobenus rosmarus*)	560	330	343	0.96	4.7	4.7	1.00	Late	0.98	9.0	8.8
Beluga (*Dalphinapterus leucas*)	675	365	360	1.01	2.3	5.0	0.46	Fairly late	0.74	9.6	7.1
Gray whale (*Eschrichtius robustus*)	30500	395	970	0.41	4.5	15.0	0.30	Fairly late	0.35	37.9	13.4
Blue whale (*Balaenoptera musculus*)	93869	335	1300	0.26	4.4	20.8	0.21	Fairly late	0.23	56.7	13.3

[a]Gestation (days) = $66.2\ M^{0.26}$.
[b]Reproductive maturity (years) = $0.75\ M^{0.29}$.
[c]$t_{2Mb} = 0.92\ M^{0.36}$.

The 1972 United Nations Conference on the Human Environment in Stockholm recommended a 10-year moratorium on the commercial killing of whales. Even this inadequate measure was rejected by the International Whaling Commission. At about that time (1974) the population of the blue whale (*Balaenoptera musculus*) was estimated to be only 6% of the virgin stock (Ehrlich et al. 1977, p. 358, 361). What could 10 years of protection do to restore this population? A recent improved allometric equation for r_{max} in mammals by Caughley and Krebs (1983) can be used to estimate the minimum time for population doubling (t_{2N_0}; see Lindstedt and Calder 1981):

$$(11) \quad t_{2N_0} = (1/1.5M^{-0.36})(\ln 4) = 0.92M^{0.36}$$

This would extrapolate to 57 years for a 93869 kg blue whale. Assuming that within a species the physiological time scale is proportionately advanced or retarded from the typical (predicted) pattern in all functions simultaneously, a correction factor for the adaptations of a species can be derived. The gestation period of the blue whale is 26% of the allometric prediction and sexual maturation takes a parallel 21% of prediction. If the assumption of other parallels in physiological timing is correct, the (average of) 23.5% of the 57-year t_{2N_0} would be 13.4 years. This is only the *minimum* time for a population doubling, to 12% of the original population, so it would probably take longer.

The Pelagic Marine Biomass

The biomasses of individual species of herbivorous mammals are apparently based upon metabolism by virtue of the scaling equilibrium between densities ($\rho \propto M^{-3/4}$) and individual metabolic rates ($\dot{E}_{st} \propto M^{3/4}$), and between total biomass ($\rho M \propto M^{1/4}$) and turnover frequency of standing crop biomass (productivity/biomass $\propto M^{-1/4}$), the productivity being directly proportioned to metabolism through the size-independence of conversion efficiency. Can this seemingly straight-forward framework be extended to apply to the vastness and complexity of an entire pelagic ecosystem? The first step in answering this question is to examine scaling exponents. Close similarity could be due to (a) the existence of general principles with wide, perhaps universal applicability or (b) coincidence or spurious correlation. If the exponents differ, there may be (c) a different set of principles operating, (d) neglect of a factor in the product of several allometric terms, (e) the crudity of small or biased samples, and/or (f) errors in calculations.

From metabolic considerations, Platt and Denman (1978) generated a theoretical unnormalized biomass spectrum for pelagic marine communities in which biomass scales as $M^{-0.22}$. This scaling is indistinguishable from a plot of empirical data for microplankton in the top 20 m of The North Pacific Central Gyre (Platt et al. 1984). The data used for this spanned six orders of magnitude in organismic mass. Clearly the theoretical ($M^{-0.22}$) and empirical scalings ($M^{-0.23}$) agree well!

However, on first inspection, the terrestrial mammals and pelagic organisms appear to be operating according to different principles. Platt and Denman (1978) noted the similarity of their theoretical plot to open ocean examples given by Sheldon et al. (1972, 1973). Sheldon et al. plotted particle concentration as a function of particle diameter on a logarithmic scale for an equatorial Pacific food chain of phytoplankton–zooplankton–micronekton–tuna and for an Antarctic food chain of phytoplankton–krill–whales. These food chains would incorporate a series of losses of energy and biomass between trophic levels that bear an inverse relationship to the steps in predator–prey size ratios.

Allometric slopes for the plots of Sheldon et al. may be estimated as the quotient of the product of mean production efficiencies (P/A, %) over the range in particle diameters.

Then we can assume that these diameters (d) scale with geometric similarity ($M \propto d^3$):

(12) $(\Delta\text{biomass})/(\Delta\text{body mass}) = [(P/A)_1 \times (P/A)_2 \times \ldots (P/A)_n]/(\Delta \log d)$ (3)

Humphrey (1979) gave P/A values which I will substitute in the above: 25.0% for "non-insect invertebrates", 9.77% for "fish and social insects," and 3.14% for "other mammals." The scalings obtained were $M^{-0.17}$ in the equatorial Pacific and $M^{-0.13}$ in the Antarctic system, not quite the $M^{-0.22}$ of Platt and Denman (1978) but within reach, considering the assumptions!

If this approach is valid, the difference in signs as well as the appropriate magnitudes of the exponents for biomass-scaling can be explained. The losses or inefficiencies of conversion of biomass between trophic-levels result in a negative allometry that masks any positive allometry of biomass among the species of a single trophic level. It remains to be demonstrated directly that the $M^{1.4}$ biomass scaling common to the mammalian herbivores can be generalized as a universal phenomenon. However, the negative biomass scaling of pelagic ecosystems cannot be taken to rule out a positive allometry within its components because of the complexity identified here. In addition the conversion efficiencies used here were actually derived on an energetic basis rather than from mass, and there is a tendency for the energy content per unit mass to increase going up the food chain because of higher fat contents.

Further argument for a common observance of the same allometric rules by both pelagic ecosystems and terrestrial mammals can be derived from the relationships between doubling times and particle diameters (d) given by Sheldon et al. (fig. 13 in Sheldon et al. 1977; fig. 9 in Sheldon et al. 1973). Applying an eye-fit to data in the former and the fits indicated by the authors in the latter, and assuming geometric similarity ($d \propto M^{1/3}$), the scalings appear to be $M^{0.23}$, to $M^{0.24}$, similar to the $M^{1.4}$ metabolic and growth time scalings noted for mammals and birds (eq. 4, discussion following eq. 9). The rate of growth or productivity would be the mass increment divided by the time:

(13) growth rate $= \Delta M/t \propto (M^{1.0}/M^{0.23}) \propto M^{0.77}$

This scaling of growth rate falls well within the range of exponents exhibited by higher vertebrates (see Case 1978; Calder 1984) and is compatible with the widespread mass-independence of productive efficiency, $M^{3/4}$ metabolic scaling, and productivity/biomass ratios proportional to $M^{-1/4}$.

Thus the advice of Steele (1974) was good: "comparison with relevant studies in land" has shown that metabolic allometry can be used as a unifying theme in the study of marine ecosystems. Kleiber would have made a good marine biologist!

Acknowledgements

A visiting professorship at the University of New South Wales and sabbatical leave from the University of Arizona provided space, library facilities, support, and time for preparation of this paper. I am grateful to Prof. T. J. Dawson and Dr. R. J. MacIntyre, U.N.S.W., for stimulating discussions and encouragement.

References

ADOLPH, E. F. 1949. Quantitative relations in the physiological constitutions of mammals. Science 109: 579–585.

BANSE, K. 1979. On weight dependence of net growth efficiency and specific rates among field populations of invertebrates. Oecologia (Berl.) 38: 111–126.

1982. Mass-scaled rates of respiration and intrinsic growth in very small invertebrates. Mar. Ecol. Prog. Ser. 9: 281–297.

BANSE, K., AND S. MOSHER. 1980. Adult body mass and annual production/biomass relationships of field populations. Ecol. Monogr. 50: 355–379.

BARTHOLOMEW, G. A. 1982. Energy metabolish, p. 57–110. *In* M. S. Gordon [ed.] Animal physiology: principles and adaptations. MacMillan, New York, NY.

BEDDINGTON, J. R., AND R. M. MAY. 1982. The harvesting of interacting species in a natural ecosystem. Sci. Am. 247(5): 62–69.

BERTALANFFY, L. V. 1957. Quantitative laws in metabolism and growth. Q. Rev. Biol. 32: 217–231.

BRODY, S. 1945. Bioenergetics and growth (reprinted 1968). Hafner, New York, NY. 1023 p.

BRYDEN, M. M. 1977. Growth and development of marine mammals, p. 1–79. *In* R. J. Morrison [ed.] Functional anatomy of marine mammals. Academic, London, New York.

CALDER, W. A., III. 1982. A tradeoff between space and time: dimensional constants in mammalian ecology. J. Theor. Biol. 98: 393–400.

1983a. An allometric approach to population cycles of mammals. J. Theor. Biol. 100: 275–282.

1983b. Body size, mortality, and longevity. J. Theor. Biol. 102: 135–144.

1983c. Ecological scaling: mammals and birds. Ann. Rev. Ecol. Syst. 14: 213–230.

1984. Size, function and life history. Harvard Univ. Press, Cambridge, MA. 431 p.

CALDER, W. A., III, AND E. J. BRAUN. 1983. Scaling osmotic regulation and body size in mammals and birds. Am. J. Physiol. 244 (RICP 13): R601–606.

CASE, T. S. 1978. On the evolution and adaptive significance of postnatal growth rates in the terrestrial vertebrates. Q. Rev. Biol. 53: 243–282.

CAUGHLEY, G., AND C. J. KREBS. 1983. Are big mammals simply little mammals writ large? Oecologia (Berl.) 59: 7–17.

DAMUTH, J. 1981a. Population density and body size in mammals. Nature 290: 699–700.

1981b. Home range, home range overlap, and species energy use among herbivorous mammals. Biol. J. Linn. Soc. 15: 185–193.

1982. Analysis of the preservation of community structure in assemblages of fossil mammals. Paleobiology 8: 434–446.

EHRLICH, P. R., A. M. EHRLICH, AND J. P. HOLDREN. 1977. Ecoscience. W. M. Freeman, San Francisco, CA. 1052 p.

FELDMAN, H. A., AND T. MCMAHON. 1983. The 3/4 exponent for energy metabolism is not a statistical artifact. Respir. Physiol. 52: 149–163.

FENCHEL, T. 1974. Intrinsic rate of natural increase: the relationship with body size. Oecologia (Berl.) 14: 317–326.

GASKIN, D. E. 1982. The ecology of whales and dolphins. Heinemann, London. 459 p.

GREZE, V. N. 1978. 3. Production in animal populations, p. 89–114. O. Kinne [ed.] *In* Marine ecology. Vol. 4. Wiley, New York, NY.

GARLAND, T. 1983a. Scaling maximum running speed and body mass in terrestrial mammals. J. Zool. (London) 199: 157–170.

1983b. Scaling the ecological cost of transport to body mass in terrestrial mammals. Am. Nat. 121: 571–587.

HEMMINGSEN, A. M. 1960. Energy metabolism as related to body size and respiratory surfaces, and its evolution. Reports of the Steno Memorial Hospital and the Nordisk Insulin Laboratorium. Vol. IX, Part II, p. 7–110.

HENNEMAN, W. W. 1983. Relationship among body mass, metabolic rate and the intrinsic rate of natural increase in mammals. Oecologia (Berl.) 56: 104–108.

HUMPHREYS, W. F. 1979. Production and respiration in animal populations. J. Anim. Ecol. 48: 427–453.

HUXLEY, J. S. 1932. Problems of relative growth (1972 reprint). Dover, New York, NY. 312 p.

KANWISHER, J. 1977. Temperature regulation, Chap. 10, p. 477–514. *In* L. Goldstein [ed.] Introduction to comparative physiology. Holt, Rinehart and Winston, New York, NY. 533 p.

KANWISHER, J., AND G. SUNDNES. 1965. Physiology of a small cetacean. p. 45–53. *In* Essays in marine physiology. Universitetets Institutt for Marin Biolog: and Statens Institutt for Hvalforskning. Hvalradets Skrifter: Scientific Results of Marine Biological Research, Nr. 48.

KING, J. R. 1974. Energetics of reproduction in birds, p. 78–107. *In* D. S. Farner [ed.] Breeding biology of birds. Natl. Acad. Sci., Washington.

KLEIBER, M. 1961. The fire of life. Wiley, New York, NY. 453 p.

LAVIGNE, D. M. 1982. Similarity in energy budgets of animal populations. J. Anim. Ecol. 51: 195–206.

LINDSTEDT, S. L., AND W. A. CALDER, III. 1981. Body size, physiological time, and longevity of homeothermic animals. Q. Rev. Biol. 56: 1–16.

MAY, R. M. 1979. Production and respiration in animal communities. Nature 282: 443–444.

PETERS, R. H. 1983. The ecological implications of body size. Cambridge Univ. Press, Cambridge, MA. 329 p.

PETERSON, R. O., R. E. PAGE, AND K. M. DODGE. 1984. Wolves, moose, and the allometry of population cycles. Science 224: 1350–1352.

PLATT, T., AND K. DENMAN. 1978. The structure of pelagic marine ecosystems. Rapp. P.-v. Reun. Cons. Int. Explor. Mer 173: 60–65.

PLATT, T., K. H. MANN, AND R. E. ULANOWICZ. 1981. Mathematical models in biological oceanography. UNESCO Press, Paris. 156 p.

PLATT, T., AND W. SILVERT. 1981. Ecology, physiology, allometry, and dimensionality. J. Theor. Biol. 93: 855–860.

PLATT, T., M. LEWIS, AND R. GEIDER. 1984. Thermo-

dynamics of the pelagic ecosystem: elementary closure conditions for biological production in the open ocean, p. 49–84. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum, London.

PRANGE, H. D., J. F. ANDERSON, AND H. RAHN. 1979. Scaling of skeletal mass to body mass in birds and mammals. Am. Nat. 113: 103–122.

RAY, G. C. 1981. 13. The role of large organisms, p. 397–413. *In* A. R. Longhurst [ed.] Analysis of marine ecosystems. Academic, London, New York. 741 p.

SACHER, G. A. 1959. Relation of lifespan to brain weight and body weight in mammals. Ciba Found. Colloq. 5: 115–141.

SCHMIDT-NIELSEN, K. 1972. Locomotion: energy cost of swimming, flying, and running. Science 177: 222–228.

SCHMITZ, O. J., AND D. M. LAVIGNE. 1984. Intrinsic rate of increase, body size, and specific metabolic rate in marine mammals. Oecologia (Berl.) 62: 305–309.

SCHNEIDER, D., AND G. L. HUNT. 1982. Carbon flux to seabirds in waters with different mixing regimes in the southeastern Bering Sea. Mar. Biol. 67: 337–344.

SHELDON, R. W., A. PRAKASH, AND W. M. SUTCLIFFE, JR. 1972. The size distribution of particles in the ocean. Limnol. Oceanogr. 17: 327–340.

SHELDON, R. W., W. H. SUTCLIFFE, JR., AND A. PRAKASH. 1973. The production of particles in the surface waters of the ocean with particular reference to the Sargasso Sea. Limnol. Oceanogr. 18: 719–733.

SLIJPER, E. J. 1979. Whales. Hutchinson, London. 511 p.

SMITH, A. H., AND N. PACE. 1971. Differential component and organ size relationships among whales. Environ. Physiol. 1: 122–136.

STAHL, W. R. 1962. Similarity and dimensional methods in biology. Science 137: 205–212.

STEELE, J. H. 1974. The structure of marine ecosystems. Harvard Univ. Press, Cambridge, MA. 128 p.

STEELE, J. H., AND E. W. HENDERSON. 1984. Modelling long-term fluctuations in fish stocks. Science 224: 985–986.

STRATHMANN, R. R., AND M. F. STRATHMANN. 1982. The relationship between adult size and brooding in marine invertebrates. Am. Nat. 119: 91–101.

TAYLOR, C. R., N. C. HEGLUND, AND G. M. O. MALOIY. 1982. Energetics and mechanics of terrestrial locomotion. I: Metabolic-energy consumption as a function of speed and body size in birds and mammals. J. Exp. Biol. 97: 1–21.

TUCKER, V. A. 1970. Energetic cost of locomotion in animals. Comp. Biochem. Physiol. 34: 841–846.

WALKER, E. P. 1975. Mammals of the world. 3rd ed. Johns Hopkins Press, Baltimore, MD.

WEIBEL, E. R. 1982. The pathway for oxygen: lung to mitochondria, p. 31–47. *In* C. R. Taylor, K. Johansen, and L. Bolis [ed.] A Companion to Animal Physiology. Cambridge Univ. Press, Cambridge, MA.

WESTERN, D. 1979. Size, life history, and ecology in mammals. Afr. J. Ecol. 17: 185–204.

1980. Linking the ecology of past and present mammal communities. Chap. 3, 41–54. *In* A. K. Behrensmeyer and A. P. Hill. [ed.] Fossils in the making: vertebrate taphonomy and paleoecology. Univ. of Chicago Press, Chicago, IL.

WESTERN, D., AND J. SSEMAKULA. 1982. Life history patterns in birds and mammals and their evolutionary interpretation. Oecologia (Berl.) 54: 281–290.

The Trophic Continuum in Marine Ecosystems: Structure and Equations for a Predictive Model

S. H. Cousins

Energy Research Group, The Open University, Milton Keynes MK7 6AA, U.K.

Introduction

The population dynamics and consequently the trophic energetics of simple two and three species systems of predators and prey or parasites and hosts are now considered well established, while the central problem of how very many species interact in a food web remains open (May 1979). This problem has also been expressed in terms of the need to shift from an understanding of single species fisheries to multispecies fisheries (Mercer 1982). Here the development of a multispecies model is first put in the context of an existing model of large ecosystems.

The traditional approach to the study of large, many species, ecosystems has been, since Lindeman (1942), to invoke a particular abstraction, the trophic level, with which to simplify ecosystem food web interactions. An organism would in the trophic level model be apportioned either wholly or in part to a particular trophic level. The green plant is easily identified as trophic level one, organisms which feed on the green plant are assigned to trophic level two and carnivores preying upon these herbivores are placed at trophic level three and so on. It is worth noting at this point that there is only one taxonomic distinction incorporated into this model. The plant kingdom excepting fungi is limited to level one and animal kingdom to trophic levels two and above. Within these kingdoms the trophic level model is truely ataxonomic in that the species identity or even the phylogenic identity of the organism is subsumed within the operational measure of trophic level content, biomass. Thus all plant species and plant parts are subsumed as plant biomass, all species of herbivores are identified as herbivore biomass and so on.

It is significant that Lindeman believed these categories to be operationally important to the evolution of ecosystems, that is, the efficiency of transfer of energy between trophic levels would become greater the higher the trophic level and that the whole chain of levels would become more efficient as a result of ecological succession. The attraction of these hypotheses was such that Lindeman's methodology (Lindeman 1942) was adopted as the basis for the International Biological Programme (IBP) 1964–74. However by the close of IBP it was evident that trophic levels could not be identified at levels three and above (Heal and Maclean 1975) and consequently Lindeman's hypotheses were not testable for large natural ecosystems.

Trophic Level: A Non-Predictive Concept

Predictive science is particularly important for ecological modelling. Prediction may not always be possible and the outcome of quite simple system interactions can be chaotic (Lorenz 1982). Peters (1980) argues that trophic level concepts are not predictive and more recently Peters (1983) that allometric models are far more predictive in ecology.

From Fig. 1 we may show formally why the trophic level concept is non-predictive. Consider trophic levels two and three and the flow between them as F_2. If the trophic level model is to be predictive then the flow, F_2, is some function, f, of the biomass at trophic levels two, M_2, and three, M_3. By definition the function f must define the flow F_2 uniquely for each combination of values for M_2 and M_3.

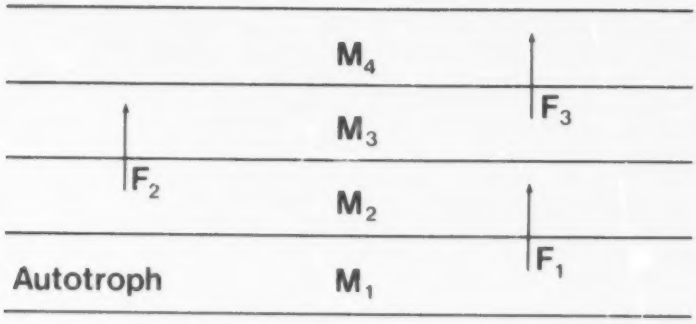


FIG. 1. Trophic level representation.

(1) $F_2 = f(M_2, M_3)$

When the flow F_2 is uniquely defined then equation (1) is predictive and consequently the trophic level model can be taken to be predictive.

We may note that a biomass M is made up of the sum of number of individuals, N multiplied by their body weight, W_i.

(2) $M = \Sigma N_i \cdot W_i$

which in continuous form is given by

(3) $m(n, w)$

where

(4) $m = \int_0^* nw \cdot dw$

Now, taking the example where M_2 is comprised of zooplankton and M_3 is made up of filter feeding fish, a flow F_2 is established. Suppose as a result of migration these fish are replaced by an equal biomass of large fish which have large mouth parts which do not allow them to feed on zooplankton yet these fish have been feeding, prior to migration, on food at trophic level 2. In the first instance F_2 will have a positive value and in the second it will be zero. Thus the flow F_2 is not uniquely defined and the hypothesis that f is a function in equation (1), is therefore rejected. However, if the data were left in the form, $m(n, w)$ for each level, then F_2 may be defined as,

(5) $F_2 = f(n_2, w_2, n_3, w_3)$

Here F_2 is uniquely determined in this instance since the flows are dependent on the number and size of particles at both levels and the test as to whether f is a function is not rejected. To make a predictive model in this form involves adding more variables to equation (5) in situations where f fails as a function. The number of variables required and the complexity of the function f will depend on the nature of the prediction being made.

Before ending the discussion of equation (5) it is useful to note that n is a quantitative variable and that, w, is a qualitative variable in the context of ecosystems. Thus to define M_2 as a resource to M_3 the quantity, n_2, at each quality, w_2, needs to be identified. Similarly to define M_3 as a predatory "force" on M_2 the quantity, n_3, of organisms at each quality, w_3, needs to be identified.

To hypothesize that transfer efficiencies change with trophic level status requires that trophic levels have objective properties with regard to the flows between levels. What equation (5) shows in contrast to equation (1) is that the flow F_2 is independent of the trophic level biomass and so trophic levels cannot possess the objective properties Lindeman envisaged.

Resource Descriptions of Ecosystems

All biological particles in the ecosystem are resources to heterotrophs, which are themselves also a resource. Although all biological particles are resources they differ in their requirements for resources. Thus autotrophs require sunlight for photosynthesis, while detritus requires no energy input and heterotrophs are dependent on ingesting or absorbing particles or dissolved organic matter whose origins are either other heterotrophs, autotrophs or detritus. These latter three categories are considered (Cousins 1980) to be fundamentally distinct trophic categories; they have different properties with respect to time and so are important to distinguish in a predictive model. With time, live organisms grow, feed and reproduce. Detritus will also change its resource status with time. Detritus may be treated as a part of the system which tends to disorder while live organisms maintain their order at the expense of greater disorder in their environment. The minimum set of categories used in the continuum model, autotroph, heterotroph and detritus are also those identified as functional groups (O'Neill 1976; Silvert and Platt 1980).

As we saw earlier, when a heterotroph feeds, the total ecosystem biomass, or biomass at a trophic level is not of interest to that organism, but only the amount of food available to that organism to ingest. For predation, which is where large heterotrophs eat smaller heterotrophs, the prey as a resource to the predator, is limited to the quantity of prey at the size of prey eaten. We may model this process in many species ecosystems as a transfer of energy or individuals, between size classes of organisms. The use of size classes to model feeding relationships in large ecosystems was first used by Elton (1927) and has been recently revived by Platt and Denman (1977) for pelagic systems and by Cousins (1978, 1980) principally for terrestrial systems.

A few simple rules appear to characterise the interactions of food webs when they are analysed by the use of size classes and each rule is discernable in Elton's original size-based description of food web operation. A comparison of these simple rules for terrestrial and pelagic food webs is given in Table 1. Making a predictive model based on

TABLE 1. Simple food web rules and processes.

Rule	Terrestrial*	Pelagic*	Process
1) Large heterotrophs eat smaller particles	Y	Y	Carnivory Herbivory Detritivory
2) Particles change size with time	Y	Y	Growth Decay
3) Some small particles eat large particles	Y	Y	Parasitism Detritivory
4) Non-particulate plants	Y	N	Herbivory
5) Non-particulate detritus	Y	N	Detritivory

*Y = Rule applies; N = Rule does not apply.

these rules therefore requires modelling the processes which are identified as corresponding to each "rule".

As is shown in Table 1 the main difference between terrestrial and aquatic plants is due to the types of green plant in the two systems. In the pelagic case the autotrophs are particulate and equations developed for predation where the heterotrophs are also treated as being particulate can be extended to apply to herbivory. Similarly particulate autotrophs and heterotrophs generate particulate detritus which in total can be viewed as a single particle size spectrum (Silvert and Platt 1980) with material and energy flowing from the small to the large particles. Thus an allometric or species size approach will apply to the whole spectrum including particulate detritus. This is not so in terrestrial systems.

The size of the terrestrial plant is much more variable than pelagic phytoplankton, and more importantly the size *per se* does not constitute an appropriate resource description for heterotrophs feeding on autotrophs again in contrast to the pelagic system. The digestibility of different terrestrial plant parts is very varied as exemplified by the extremes of wood and seeds. Thus while number of phytoplankton particles, n_A, of particular weights, w, can describe the quantity of autotroph resources of different qualities, for the terrestrial plant the quality of the resource is described by v, the digestibility (see below) of a particular plant material and the quantity or the mass of that material present.

Figure 2 shows a comparison of the resource descriptions of the mass distribution of terrestrial autotroph material and the number distribution of pelagic autotrophs. The digestibility variable, v, has the range of values 0 to 1 corresponding to 0 to 100% assimilability of the plant material. The broad categories of plant parts, wood, leaves, and seeds are also important functional categories which have interesting temporal behaviours (Parkin and Cousins 1981). The range of weights of phytoplankton is given as 0 to w' the largest phytoplankter in the system.

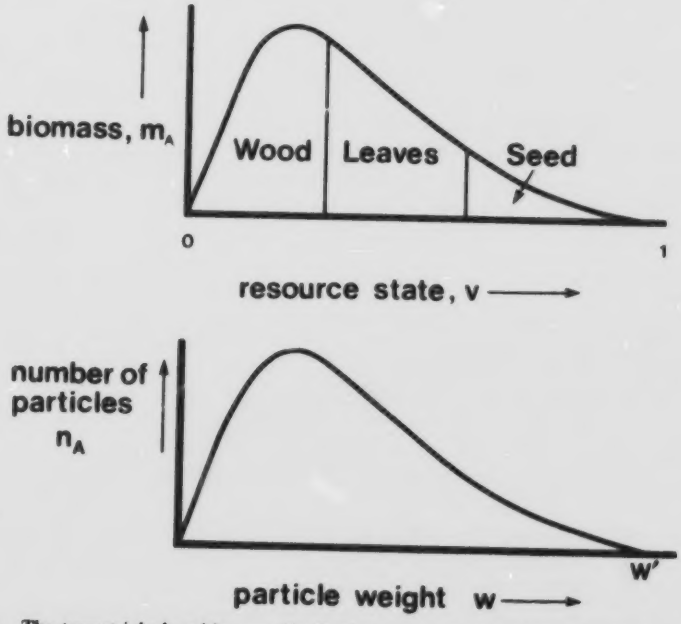


FIG. 2. The terrestrial plant biomass distribution and a phytoplankton number distribution.

The Trophic Continuum Model

The three fundamental categories of the trophic continuum, autotrophs, heterotrophs, and detritus are those identified above as constituting distinct groups in their requirements for resources and as being distinct in the way that they change resource state with time. Figure 3 represents the trophic continuum as a cylindrical graph which relates the three fundamental categories and also represents changes in resource state occuring within the categories themselves. Row *A* represents the resource states of the many plant species and plant parts which are available for ingestion by heterotrophs. State A_{m-1} represents all primary products of photosynthesis in the system. States A_m to A_n represent the different chemical and structural forms of plant biomass and correspond to *v* states or the number distribution of marine autotrophs shown in Fig. 2. Translocation and chemical transformation of the products of photosynthesis create the transfer of material between resource states in the plant. This contrasts to the Lindeman model where biomass only changes state when something is eaten.

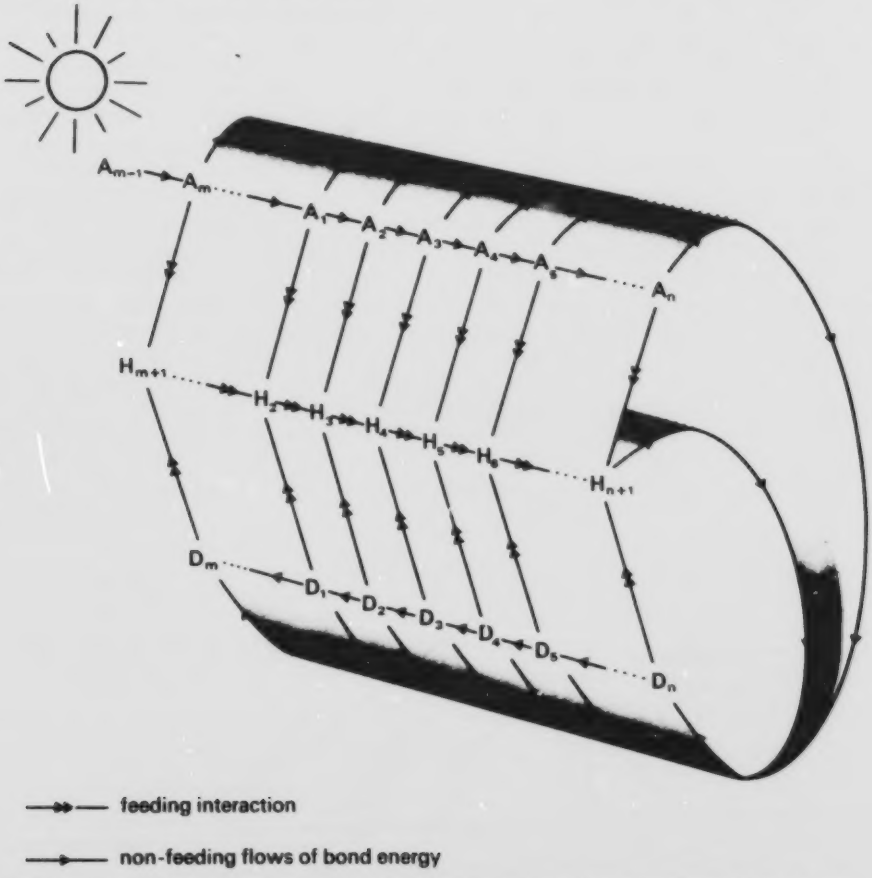


FIG. 3. The trophic continuum showing translocation, chemical transformation within the autotroph, heterotroph interactions, and detritus fragmentation.

The autotrophs are fed upon by heterotrophs of sizes H_m (small) to H_n (large). Feeding interactions (carnivory) also occur between H states. With time individuals will grow and will therefore change their size category, and therefore alter their resource state in the model which is again in contrast to Lindeman where, of itself, growth does not affect trophic status.

Detritus is also fed upon by heterotrophs and is also generated from both the autotroph and heterotroph compartments of the model. Detritus in the form of carcasses, dung and plant litter as well as exudates and exuviae constitute the variety of states D_m to D_n. With time each item of detritus changes its resource state towards the state of greatest dispersal, D_m. In the pelagic system D_m represents dissolved organic matter and D_n the newly dead carcass of the largest heterotroph, H_n in the system. The trophic continuum shown in Fig. 3 is simplified in that only four routes for the transfer of energy are shown at any single state.

Marine Systems

The particulate autotrophs and heterotrophs and their particulate litter, carcasses and dung allow each of the autotroph, heterotroph, and detritus resource states of the trophic continuum to be represented by a number distribution of particles of different weights. The autotroph number distribution, $n_A(w)$, has already been illustrated in Fig. 2b. Similar distributions for heterotrophs, $n_H(w)$, and detritus, $n_D(w)$ can be defined for the number of heterotroph and detritus particles of different weights, w.

Figure 4 shows the interactions between these distributions. The figure may be considered as an end view of the cylindrical representation of the trophic continuum of Fig. 3. The three number distributions of particle weight are the resource state distributions of the continuum.

Predictive Models

The processes by which the particle distributions undergo change with time have been identified in Table 1. These processes occur as an interaction between distributions, together with carnivory, which occurs within the heterotroph number distribution, are shown in Fig. 4; the remaining processes of growth and decay also occur within the distributions but are not shown. Modelling these processes provides a predictive model of the change in the number distributions of autotrophs, heterotrophs, and detritus particle weights.

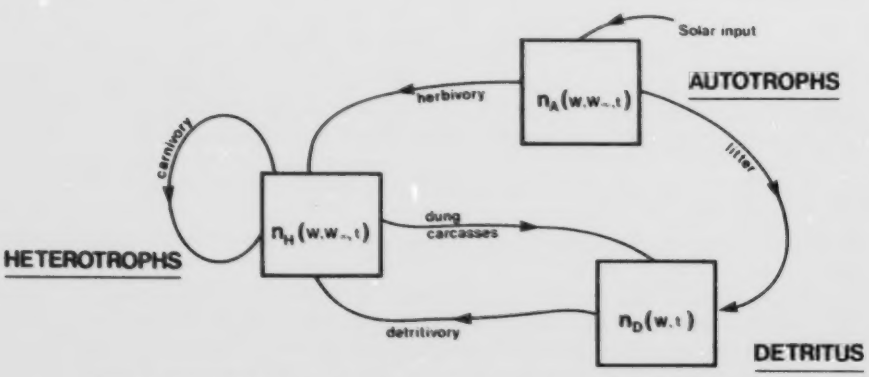


FIG. 4. Inputs and outputs to the functional groups of the trophic continuum.

There is a wide literature relating body size to the processes required for a predictive model. Peters (1983) has assembled over 1000 allometric relationships for various parameters and has reviewed their importance for ecology. He sees these relationships as providing a basis for predictive ecology. The substance of this claim is that by knowing an organism's body weight, predictions can be made about the processes important to its ecology. These include metabolic rate, reproductive rate, mobility, daily food demand, size of food particle eaten, and so on. However, depending on the degree of accuracy required these allometric data may require splitting into the appropriate taxonomic groups such as phyla.

Here we will consider two of the processes in more detail, growth and heterotrophy and also discuss the role of taxonomic groupings in the trophic continuum.

Growth

Knowing the present state of an ecosystem in terms of the size distribution of particles is useful but needs to be further defined if we are to predict the size distribution at some future time, t. As noted in Table 1, particles change size, notably live particles grow. Two particles may be the same size at time t_0 but be very different sizes at t_1 because one particle is adult at time t_0 with little prospect of growth and the other is a juvenile on a growth path to being a much larger organism. Thus we need to define the number distribution of particles, n, as a function of the particles, present weights, w, and the asymptotic weights, w_∞ to which they are growing.

$$n(w, w_\infty, t) \tag{6}$$

It is of interest that w_∞ is, under certain circumstances, the species identity of the particle as is discussed in the taxonomy section below. However, the consideration of the w_∞ variable does not require the identification of species in the trophic continuum (Cousins 1980) or the biomass spectrum (Platt and Denman 1977). This is because we may make the assumption that there is a continuum of asymptotic species weights between the smallest and largest particle in the ecosystem, just as Platt assumes a continuous distribution of individuals. The w_∞ variable is vital to the behaviour of the model since it prevents small particles, e.g., bacteria, growing into whales. We might also wish to indicate where species are absent from the continuum of asymptotic adult weights. The inclusion of the adult weight variable w_∞ allows investigation of the phenomenon of changing diet with changing size which was termed metaphoetesis by Hutchinson (1959).

Phytoplankton Population Growth

Here we model the growth of a population of phytoplankton which reproduces by binary fission. The growth trajectory is between w_1 which is half the weight of w_2, and w_2, the weight at which binary fission occurs. Assuming a mathematically simple growth curve, then the growth of an individual after time, t, is given by

$$w = w_\infty(1 - e^{-t/\tau}) \tag{7}$$

where w_∞ and the constant τ can be determined from empirical values for w_1, w_2 and time taken to grow from w_1 to w_2.

Equation (7) indicates that individuals grow relatively rapidly away from their initial size when fission has just taken place and then grow progressively more slowly as the point of fission is reached again. This has interesting consequences for the detailed shape of the biomass spectrum or number distribution $n(w, w_\infty, t)$. For any one species individuals spend more time and are therefore more frequently observed in the larger positions of their

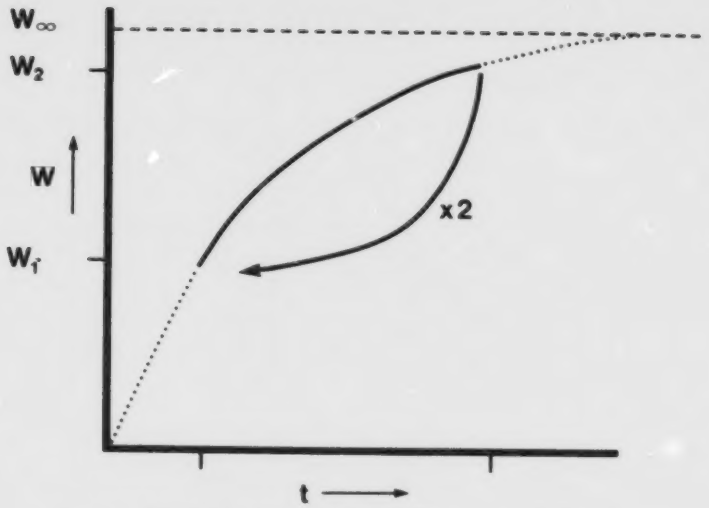


FIG. 5. Growth and binary fission of an individual phytoplankter.

species growth curve. This is in opposition to the general decreasing trend in number of individuals with increasing particle weight first observed by Elton (1927). The resolution of these two processes should produce a "serrated" biomass distribution with exponential growth away from a perfectly flat distribution of logarithmic weight classes.

The creation of a "serrated" particle size distribution is demonstrated by Cousins and Parkin (unpublished data) for a multispecies model of phytoplankton undergoing binary fission. First let us consider the single species case.

From equation (7) for the individual the change in the phytoplankton particle weight distribution $n_A(w, t)$, is given by:

$$(8) \quad n_A(w, t) = 2\alpha e^{t/\tau} \cdot n_A(w_0, 0)$$

where

$$(9) \quad \alpha = \frac{w_\infty - w_1}{w_\infty - w_2}$$

The fact that an individual spends more time at larger weights is shown by individuals "bunching" at higher weights in the species distribution, n_A. Figure 6 shows this effect for a particular set of starting conditions. $n_A(w_0, 0)$, where a cohort of individuals is equally distributed over the size range w_1 to some size, w_2, 10% larger than w_1. Note that the cohort of individuals becomes narrower and taller prior to fission showing that the same number of individuals are found to be distributed in fewer weight classes.

By taking nine species of different asymptotic adult weights w_∞, and giving initial conditions such that the species abundances are proportional to $1/w_\infty$, then the growth of the nine species populations is shown in Fig. 7. The population at time t is given by

$$(10) \quad \sum_{i=1}^{9} n_i(w, t) = \sum_{i=1}^{9} 2\alpha_i e^{t/\tau} \cdot n_i(w_0, 0)$$

and generates the serrated curve of particle number against particle weight, $n_A(w, w_\infty, t)$. The number of serrations is a function of the number of species present although the

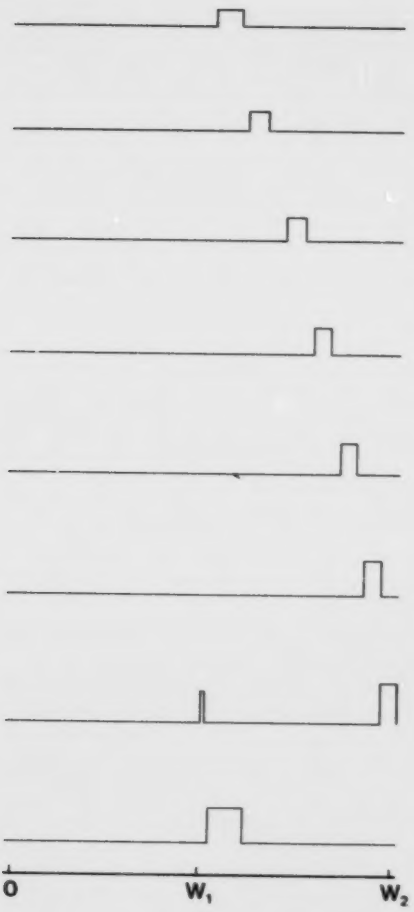


FIG. 6. A cohort of phytoplankton undergoing growth and binary fission.

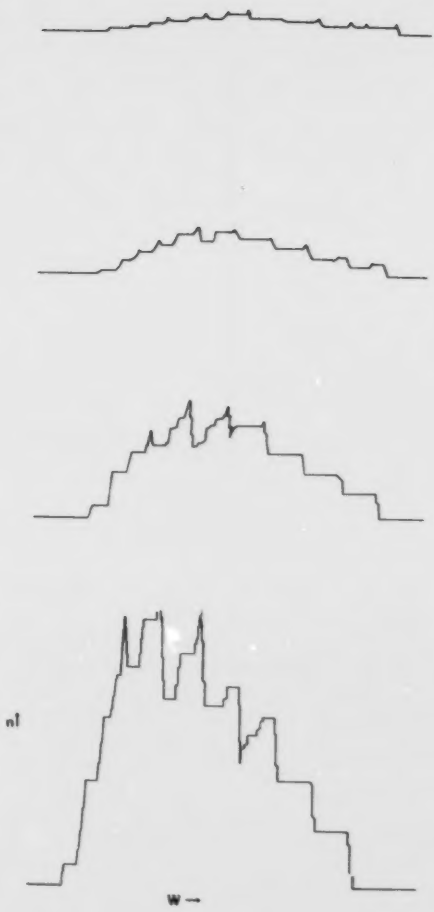


FIG. 7. Cumulative population growth of nine species of binary fissioning phytoplankters.

complex life-histories of many phytoplankton may create more than one "notch" per species. Detailed phytoplankton number distributions of particle weight do appear serrated at the micro-scale (Sheldon et al. 1972).

Predation

The study of this form of heterotrophy at the ecosystem level has until recently been limited to trophic level theory. Within this theory there is little if any attention paid to the mechanism of predation itself since the concern is to define organisms as either a level above or below one another dependent on who eats whom. However at the single species level the mechanisms of predation have been much studied and theories of foraging and food capture have been developed. These have been reviewed by Pyke et al. (1977).

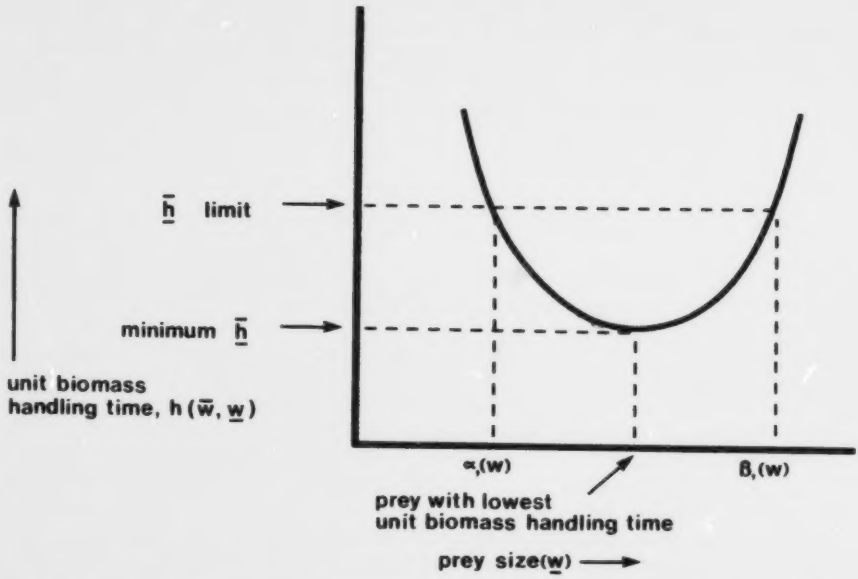


FIG. 8. The time taken for a predator to handle one unit weight of prey biomass for prey of different sizes.

Foraging theory has an allometric basis and can be extended to apply to a large ecosystem containing many species. A general foraging model for free swimming aquatic animals has been developed by Gerritsen (1984).

From optimal foraging theory heterotrophs will encounter other particles at a rate dependent on the abundance of heterotrophs of a given size and the abundance of other particles in the system. A single heterotroph will encounter particles at a rate dependent on the density of particles and the distance travelled by the ingesting organism. Only certain of those encounters will lead to ingestion. Elton (1927) observed that there is both an upper and a lower limit of prey size taken by any predator. The upper limit was determined by the power required by the predator to pursue catch and kill the prey organism while the lower limit is controlled by the inefficiency of collecting small particles even though small particles are likely to be progressively more abundant than large. These phenomena are linked by the concept of handling time which is the time taken to subdue and ingest the particle. The time spent searching for food has been called foraging time. Handling time divided by the weight of the particle gives the unit biomass handling time. This will reach a minimum for the optimal particle size ingested by a particular heterotroph and unit biomass handling times increase for either larger or smaller particles as is shown in Fig. 8. There is a limit shown as a horizontal line above which handling times reduce foraging times to such a degree that food intake is reduced as a whole. Thus, in our model the diet must be satisfied between the limits α, β as death by starvation will occur outside this range. The establishment of prey choice by the heterotroph corresponds to Rosen's (at this conference) concept of the organism's questioning the environment. In the model presented here (Parkin and Cousins 1981) the demand for food by a heterotroph, is limited

to that required for metabolism, growth and reproduction, and satiation occurs when these demands have been met. The shape of the unit biomass handling time curve, shown in Fig. 8 will vary for different feeders, for example, as between raptoral and filter feeders. All particles within the weight range of the diet are eaten if encountered irrespective of the particle abundance. Abundance of the particles, however, will affect the limits α, β of prey choice. An optimal foraging equation for the whole system is given in the Appendix.

Parasitism

Parasitism is the form of heterotrophy where small heterotrophs derive their food intake from larger organisms, which unlike predation, normally continue to live. The inclusion of parasitism in a multispecies model of large ecosystems is particularly challenging and solutions to this problem run counter to the general approach of allometric modelling of large systems. For predation, species identity can be largely ignored for both predator and prey because for a predator of given size will eat any prey species which falls within the range of food particle sizes of the predators optimal diet. Small heterotrophs cannot generally eat large heterotrophs and if they could the system would rapidly collapse due to the greater rate of increase of small organisms. Parasitism appears to occur only under the strictly limited condition that parasites tend to be host specific and are dependent on the host's (patchy) distribution. This may force a detailed spatial dimension on an otherwise temporal model.

In the face of such difficulty the tendency has been, so far, to ignore parasitism as a part of the food web dynamics of large systems. This is true for Parkin and Cousins (1981). However a basis for modelling parasite interactions lies in the pattern of distribution of species size, w_∞ found in the number distribution $n(w, w_\infty, t)$. The distribution of species size also appears to have an allometric basis (May 1978). Thus rather than treat w_∞ as having a continuum of values from the smallest to the largest creature in the system, we should expect a series of discrete values for w_∞ akin to Hutchinson's ratios for trophic apparatus (Hutchinson 1959). The importance of parasitism in the food web may be indicated by deep serrations on the particle size distribution curve since like the modest serrations caused by growth, Fig. 7, this is a species based phenomenon.

Taxonomy

Elsewhere (Cousins 1983) I have concluded that "the discovery of the importance of body size phenomena in trophic ecology is also the rediscovery of the importance of taxonomic ecology. Allometric relationships appear to hold most precisely within taxonomic groups". Thus while the aim of the trophic continuum model is to define the ecosystem with a minimum of taxonomic distinction this is controlled by the precision of the output which is required.

Fenchel (1974) has identified three taxonomic distinctions which apply to the intrinsic rate of natural increase; these are single celled organisms, heterotherms and homeotherms. The identification of the difference between heterotherms and homeotherms was also shown to be of importance to the shape of the growth curve which, through metaphoetesis, is of importance to the food web (Hutchinson 1959). Metaphoestesis is more important in heterotherms than homiotherms where parents forage for their offspring. Phillipson (1981) has considered more extensively the relationship between allometric phenomena and phylogeny. Much finer levels of taxonomic identity are relevant to parasitism. Parasites which are of considerable functional importance to the food web, operate at the species level. But as has been noted, the distribution of species size within taxa offers the potential for allometric generalisation about parasitism.

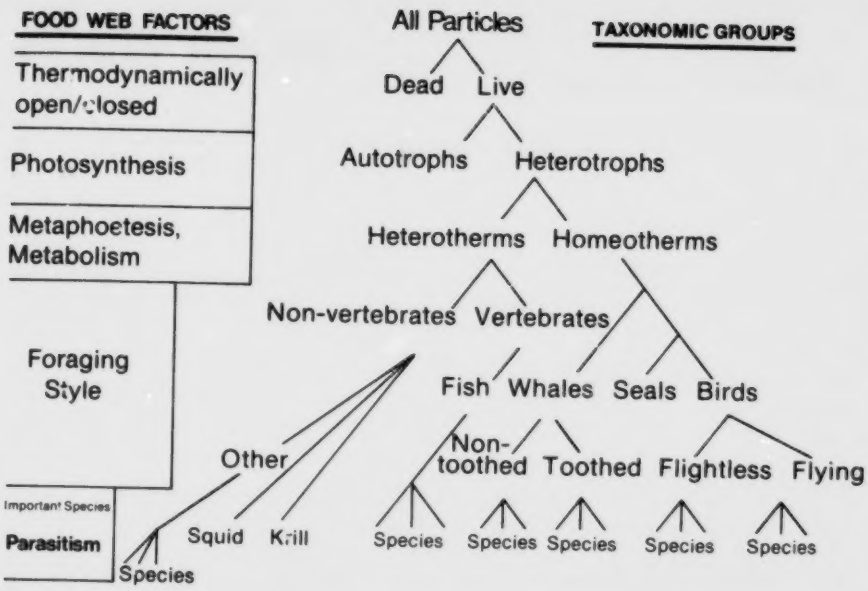


FIG. 9. Taxonomic classifications and their significance to food web factors in the Antarctic food web.

Modes of locomotion and modes of feeding can be functionally important to the operation of the food web. Thus we may wish to distinguish birds from amphibians of the same weight, or filter feeders from raptoral feeders. The degree of taxonomic distinction depends on the question that the model is required to address. Bonner (1981) is concerned with the food web relationships between bird, seal and whale populations as they have been or will be affected by change in whale and krill stocks. A set of taxonomic groups useful in this context is given in Fig. 9. The functional importance of these taxa in the trophic continuum is identified. Each taxon is represented in the trophic continuum model by a number distribution $n(w, w_\infty, t)$ and by allometric characteristics for foraging, reproduction, metabolic rates and so on. Thus the size of food taken by baleen compared to toothed whales is very different and central to the question posed since toothed whales will prey on seals and swimming birds while baleen whales will feed on krill. When baleen whales are separated out some interesting allometric properties emerge as is shown by the trend in species number and baleen diameter with increasing latitude in Fig. 10.

The number distribution of all heterotrophs in the continuum model can thus be seen as the sum of number distributions of various taxa, n_i

$$(11) \quad \Sigma\, n_i(w_i, w_\infty, t)$$

The cumulative number distribution is a symbolic representation of the predictive model of the trophic continuum which has been developed. These number distributions determine the flows between model compartments and not vice versa. Thus this is a force rather than a flow model.

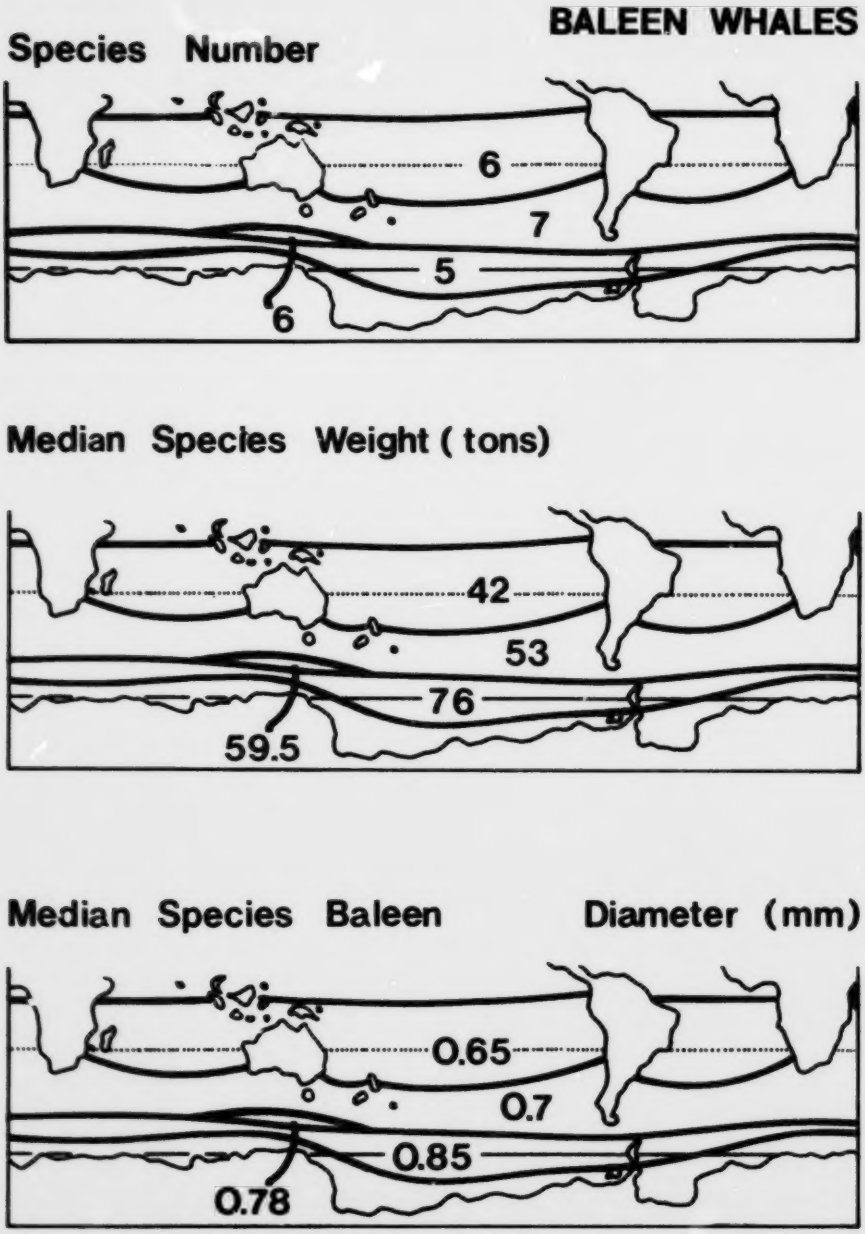


FIG. 10. The distribution of baleen whale species and their characteristics, from Watson (1981).

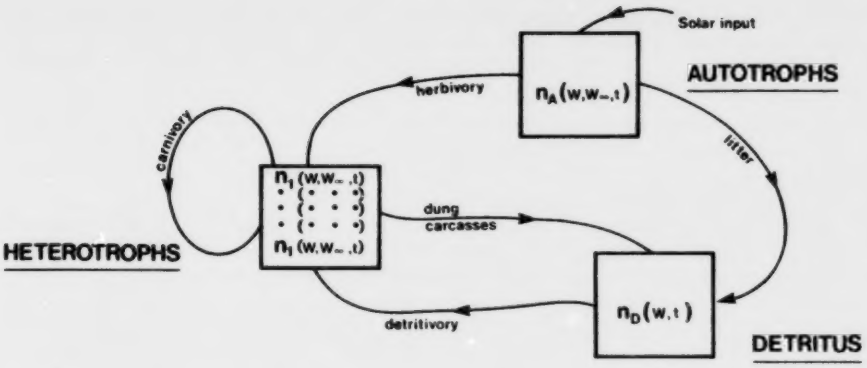


FIG. 11. Inputs and outputs of the functional groups of the trophic continuum showing heterotrophs separated into taxonomic groups.

Since autotrophs and detritus are particulate their number distributions may also be added to equation (14) to give a single particle size distribution for the purpose of modelling heterotrophy. Carnivory arises when the particles in the optimal diet are larger than phytoplankton or detritus particles, but no distinction is necessary in the structure of the model with all particle transfers to heterotrophs classed as heterotrophy, see appendix. The overall structure of the trophic continuum model is shown in Fig. 11.

Discussion

The development of models of large, multispecies ecosystems is at an early stage appropriate to our present level of understanding of the interactions of multispecies food webs. However, the trophic level concept finds no place in this new analysis. There is no implied criticism of Lindeman's work by that statement since Lindeman's lasting achievement (1942) was to write a highly stimulating paper on the nature of ecological succession using Hutchinson's notation for the trophic level. It is interesting too that Hutchinson's empirical work (1959) on the study of food webs and species assemblies (Hutchinson and MacArthur 1959) has contributed the important concepts of metaphoetesis and morphological ratios which are of importance to the allometric analysis of food web properties. But to develop these allometric models unconditionally it has been necessary to show that the concept of trophic level does not carry over into what we may now call allometric ecology.

The structure of the trophic continuum model as described in this paper has been to create a whole ecosystem model based on processes which occur at the micro-scale and which are relevant to food web interactions. The incorporation of organism growth by giving the particle two weight descriptors, its present weight and its asymptotic adult weight, and the inclusion of an optimal foraging strategy as the basis for predation are crucial features of the model developed. The w_∞ variable provides the option for incorporating species identity into the trophic continuum model if it is required. Where it is not appropriate to identify species, the w_∞ variable will still ensure that bacteria sized particles will not grow to whale size proportions with the model. When field data is collected some assumptions must be made about the relationship between the observed particle distribution, w, and w_∞ the destinations of those particles.

The realisation that heterotrophs change their diet as they grow has yet to have the radical impact that it deserves on how we perceive the operation of food webs. This process of metaphoetesis (Hutchinson 1959) upsets confidence in our understanding of how even simple two or three species food webs behave. In contrast, metaphoetesis provides a mechanism which shows the way in which multispecific interactions will occur, see also Pope and Knights (1982). Species interactions are probably of a much higher order than previously suspected because as (Gulland 1983) has noted for fish, competition for food occurs with different species at different life stages.

In spite of the complexity of the model described in this paper the model's emergent properties may be quite simple even given a wide range of initial conditions. If that is the case, simpler models may adequately describe the outcome of perturbations to any number distribution of organisms. Ulanowicz (at this conference) has defined emergent properties of ecosystems as thermodynamic properties. With the concept of ascendency, Ulanowicz (1982) is concerned, as was Lindeman, to identify the change in ecosystem state variables which occur during succession or during other forms of ecosystem evolution. It would be useful to examine Ulanowicz's concepts taking body size as the state variable and to examine the properties of the trophic continuum or biomass spectrum as they are affected by succession. Two other thermodynamic questions are of interest, the first is the effect of temperature on the operation of ecosystem processes in an allometric model, and the second concerns the $m(n, w)$ representation of biomass as a resource descriptor. Here Georgescu-Roegen (1971) stresses that different resource states, in this case values of w, are analogous to different entropy states.

Platt and Silvert (1981) have sought to explain the importance of organism size in biological processes. It is a substantial achievement that they have established dimensionalism as providing such an explanation. But further challenges exist and there are food web processes which we may call informational which are only partially explained by dimensional analysis of receptor organs Maiorana (1981), although information content is itself dimensionless. While predation may be understood in terms of the relative muscle volumes of predator and prey, parasites cannot overpower their hosts but must find some specialised technique, some species specific information, by which they can defeat the host's defenses. The predator's search for prey *in an environment* is also an information processing problem. While aspects of these information based problems may yield to allometric methods, thermodynamics may also provide analogies which are useful to increase our understanding.

Acknowledgements

I wish to thank the SCOR workshop organisers for inviting me to attend and I thank Howard Parkin for his help on the diatom model. This research was supported by The Open University.

References

BONNER, W. N. 1981. The krill problem in Antarctica. Oryx 16: 31–37.

COUSINS, S. H. 1978. Trophic models — was Elton right? 2nd Int. Cong. of ecol., Jerusalem. Energy Research Group Report 029, The Open University, Milton Keynes, U.K.

1980. A trophic continuum derived from plant structure, animal size and a detritus cascade. J. Theor. Biol. 82: 607–618.

1983. An alignment of diversity and energy models of ecosystems. Ph.D. thesis, The Open University, Milton Keynes, U.K.

ELTON, C. S. 1927. Animal ecology. Sidgwick and Jackson, London.

FENCHEL, T. 1974. Intrinsic rate of natural increase: the relationship with body size. Oecologia(Berl.) 14: 317–326.

GEORGESCU-ROEGEN, N. 1971. The entropy law and

the economic process. Harvard University Press, Cambridge MA.

GERRITSEN, J. 1984. Size efficiency reconsidered: a general foraging model for free-swimming aquatic animals. Am. Nat. 123: 450–467.

GULLAND, J. A. 1982. Why do fish numbers vary? J. Theor. Biol. 97: 69–75.

HEAL, W. O., AND S. F. MACLEAN, JR. 1975. *In* W. H. van Dobben and R. H. Lowe-McConnell [ed.], Unifying concepts in ecology, p. 89–108. Junk publications, The Hague.

HUTCHINSON, G. E. 1959. Homage to Santa Rosalia: or why are there so many kinds of animals? Am. Nat. 93: 117–125.

HUTCHINSON, G. E., AND R. H. MACARTHUR. 1959. A theoretical ecological model of size distributions among species of animals. Am. Nat. 93: 117–125.

LINDEMAN, R. L. 1942. The trophic-dynamic aspect of ecology. Ecology 23: 399–418.

LORENZ, O. E. 1982. The strange attractor theory of turbulence. Ann. Rev. Fluid. Mech. 14: 347–364.

O'NEILL, R. V. 1976. Ecosystem persistence and heterotrophic regulation. Ecology 57(6): 1244–1253.

MARIOANA, V. C. 1981. Prey selection by sight: random or economic? Am. Nat. 118: 450–451.

MAY, R. M. 1978. The dynamics and diversity of insect faunas. *In* L. A. Mound and N. Waloff [ed.], Diversity of insect faunas. Blackwell Scientific Press, Oxford.

1979. The structure and dynamics of ecological communities. *In* R. M. Anderson, B. D. Turner, and L. R. Taylor [ed.], Population dynamics. Blackwell Scientific Press, Oxford.

MERCER, M. C. [ED.] 1982. Multispecies approaches to fisheries management advice. Can. Spec. Publ. Fish. Aquat. Sci. 59.

PARKIN, H., AND S. H. COUSINS. 1981. Towards a global model of large ecosystems; equations for the trophic continuum. Energy Research Group Report 041, The Open University, Milton Keynes, U.K.

PETERS, R. H. 1980. Useful concepts for predictive ecology. Synthese 43: 257–269.

1983. The ecological implications of body size. Cambridge University Press, Cambridge.

PHILLIPSON, J. 1981. Bioenergetic options and phylogeny. *In* C. R. Townsend and P. Calow. [ed.] Physiological ecology. Blackwell Scientific, Oxford.

PLATT, T., AND K. DENMAN. 1977. The organisation of the pelagic ecosystem. Helgol. Wiss. Meeresunters 30: 575–581.

PLATT, T., AND W. SILVERT. 1981. Ecology, physiology, allometry and dimensionality. J. Theor Biol. 93: 855–860.

POPE, J. G., AND B. J. KNIGHTS. 1982. Simple models of predation in multi-age multispecies fisheries for considering the estimation of fishing mortality and its effects. *In* M. C. Mercer [ed.], Multispecies approaches to fisheries management advice. Can. Spec. Publ. Fish. Aquat. Sci. 59.

PYKE, G. H., H. R. PULLIUM, and E. L. CHARNOV. 1977. Optimal foraging: A selective review of theory and tests. Q. Rev. Biol. 52: 137–154.

SHELDON, R. W., A. PRAKASH, AND W. H. SUTCLIFFE. 1972. The size distribution of parasites in the ocean. Limnol. Oceanogr. 17: 327–340.

SILVERT, W., AND T. PLATT. 1980. Dynamic energy-flow model of particle size distribution in pelagic ecosystems, p. 754–763. *In* W. C. Kerfoot [ed.] Evolution and ecology of zooplankton communities. The University Press of New England, NH.

ULANOWICZ, R. E. 1980. An hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

WATSON, L. 1981. Sea guide to whales of the world. Hutchinson, London.

Appendix

The equations presented here form a global model of large ecosystems. The model is specified at any time by number distributions of autotrophs, heterotrophs, and detritus. Dynamic events are determined by a weight dependent appetite function, an optimal foraging strategy, growth, reproduction, and detritus decay terms.

A full description of the model and its derivation is given in Parkin and Cousins (1981).

Notation

n_A the number distribution of autotroph particles by weight

n_H the number distribution of heterotrophs by weight

n_D the number distribution of detritus particles by weight

ϕ has value 1 when predator is foraging within optimal foraging limits and 0 when outside these limits; see Fig. 8.

f the weight of food per unit time derived by a heterotroph of weight w while optimally foraging

k the requirement for food per unit time for metabolism, growth and reproduction for a heterotroph of weight, w

p predator−prey encounter rate

w the current weight of a particle

w_∞ the asymptotic adult weight of a particle

α the minimum particle size ingested to maximise food input per unit time to a heterotroph of size, w

β the maximum particle size ingested to maximise food input per unit time to a heterotroph of size, w

h the time taken to handle a unit weight of prey biomass by a heterotroph of weight, w

d death rate due to starvation

b a Gaussian distribution function to allocate births to individual weights for a parent heterotroph of weight, w

r the number of offspring per individual of reproductive age per unit time

γ fraction of the asymptotic adult weight at which reproduction begins

g a Gaussian distribution function for allocating dung

a fraction of food assimilated

C_F natural log of the inverse of the time taken for a particle to fragment to half its present weight

Bar notation: sub bar indicates variable as prey, e.g., $\underline{w}$ is the weight of a prey particle, super bar indicates variable as predator, e.g., $\bar{w}$ is the weight of the same particle acting as a predator, where both bars are used, e.g., $\underline{\bar{h}}$ this is shorthand for $h(\bar{w}, \underline{w})$.

1. The equation for the plant (autotroph) is

$$(1)\quad \frac{\partial n_A}{\partial t} = \left(\frac{\partial n_A}{\partial t}\right)_{\text{heterotrophy}} + \left(\frac{\partial n_A}{\partial t}\right)_{\text{growth}} + \left(\frac{\partial n_A}{\partial t}\right)_{\text{litter}}$$

where

$$(2)\quad \left(\frac{\partial n_A}{\partial t}\right)_{\text{heterotrophy}} = -n_A \int_0^\infty \bar{\Phi} \cdot \frac{\min\{\bar{f}, \bar{k}\}}{\bar{f}} \cdot \frac{\bar{p}\left(\int_{\bar{w}}^\infty \bar{n}\, dw_\infty\right) d\bar{w}}{1 + \int_\alpha^\beta \underline{w}\,\underline{\bar{h}}\,\bar{p}\left(\int_{\underline{w}}^\infty (\underline{n_A} + \underline{n_H} + \underline{n_D} \ldots \underline{n_i})\, dw_\infty\right) d\underline{w}}$$

$$(3)\quad \bar{f} = \frac{\int_\alpha^\beta \underline{\bar{P}}\,\underline{w} \int_{\underline{w}}^\infty (\underline{n_A} + \underline{n_H} + \underline{n_D} \ldots \underline{n_i})\, dw_\infty\, d\underline{w}}{1 + \int_\alpha^\beta \underline{w}\,\underline{\bar{h}}\,\bar{p} \int_{\underline{w}}^\infty (\underline{n_A} + \underline{n_H} + \underline{n_D} \ldots \underline{n_i})\, dw_\infty\, d\underline{w}}$$

$$(4)\quad \left(\frac{\partial n_A}{\partial t}\right)_{\text{growth}} = \frac{\partial}{\partial w}\left(\frac{w - w_\infty}{t} \cdot n_A\right)$$

$\left(\frac{\partial n_A}{\partial t}\right)_{\text{litter}}$ ~ not modelled for marine, only terrestrial plant.

II. The equation for heterotrophs is

$$(5)\quad \frac{\partial n_H}{\partial t} = \left(\frac{\partial n_H}{\partial t}\right)_{\text{heterotrophy}} + \left(\frac{\partial n_H}{\partial t}\right)_{\text{starvation}} + \left(\frac{\partial n_H}{\partial t}\right)_{\text{growth}} + \left(\frac{\partial n_H}{\partial t}\right)_{\text{reproduction}}$$

where,

$$(6)\quad \text{So, } \frac{\partial n_H}{\partial t}_{\text{heterotrophy}} = -\, n_H \int_0^{\infty} \bar{\Phi} \cdot \frac{\min\{\bar{f}, \bar{k}\}}{\bar{f}} \cdot \frac{\bar{p}_1 \left(\int_{\bar{w}}^{\infty} \bar{n}_1 \, dw_\infty \right) d\bar{w}}{1 + \int_{\alpha}^{\beta} \underline{w}\, \underline{\bar{h}}\, \underline{\bar{p}} \left(\int_{\underline{w}}^{\infty} (\underline{n}_A + \underline{n}_H + \underline{n}_D \ldots \underline{n}_i) dw \right) d\underline{w}}$$

$$(7)\quad \left(\frac{\partial n_H}{\partial t}\right)_{\text{starvation}} = -n_H d \max\left\{0, 1 - \frac{f}{k}\right\}$$

$$(8)\quad \left(\frac{\partial n_H}{\partial t}\right)_{\text{growth}} = \frac{\partial}{\partial w}\left(\frac{w - w_\infty}{t}\right) \cdot n_H\Big)$$

$$(9)\quad \left(\frac{\partial n}{\partial t}\right)_{\text{reproduction}} = br \int_{\gamma w_\infty}^{w_\infty} \bar{n}\, d\bar{w} \quad [\bar{w} \text{ is the parent not offspring}]$$

III. The equation for detritus is,

$$(10)\quad \frac{\partial n_D}{\partial t} = \left(\frac{\partial n_D}{\partial t}\right)_{\text{heterotrophy}} + \left(\frac{\partial n_D}{\partial t}\right)_{\text{carcass}} + \left(\frac{\partial n_D}{\partial t}\right)_{\text{dung}} + \left(\frac{\partial n_D}{\partial t}\right)_{\text{fragmentation}}$$

where,

$$(11)\quad \left(\frac{\partial n_D}{\partial t}\right)_{\text{heterotrophy}} = -n_D \int_0^{\infty} \bar{\Phi} \cdot \frac{\min\{\bar{f}, \bar{k}\}}{\bar{f}} \cdot \frac{\bar{p}_1 \left(\int_{w}^{\infty} \bar{n}\, dw \right) d\bar{w}}{1 + \int_{\alpha}^{\beta} \underline{w}\, \underline{\bar{h}}\, \underline{\bar{p}}\, \underline{n}_D\, dw}$$

$$(12)\quad \left(\frac{\partial n_D}{\partial t}\right)_{\text{carcass}} = \int_w^{\infty} n_H \, dw_\infty \, d \max\left\{0, 1 - \frac{f}{k}\right\} + \text{plant litter term}$$

$$(13)\quad \left(\frac{\partial m_D}{\partial t}\right)_{\text{dung}} = g(1 - a) \int_0^{\infty} \min\{f, k\} \int_w^{\infty} n_H \, dw_\infty \, dw$$

$$(14)\quad \left(\frac{\partial n_D}{\partial t}\right)_{\text{fragmentation}} = \frac{1}{w} \frac{\partial}{\partial w} (C_F w^2 n_D)$$

IV. FLOW ANALYSIS

Ecosystem Flow Analysis

BRUCE HANNON

Department of Geography, Energy Research Group & Illinois Natural History Survey, University of Illinois, Urbana, IL 61801, USA

Introduction

Long-term observation of the interaction among the components of a living system will suggest that material flows are taking place between some of the components. The system will be seen to absorb a relatively high free energy input either in the form of light, introduced organic matter, or chemical substances. The components may change in total biomass from time to time, and the whole system will give off heat and perhaps exported substances. Some of these flows can be measured directly with, for example, chemostats or tracers, or the flows might be implied from measures of the stock levels and the stock level changes of the various substances in the system.

The ecological literature contains many examples of empirically determined flows and stocks in real ecosystems. Attempts to arrange these flows in systematic form for the purpose of determining their direct and indirect relationships, apparently began with Hannon (1973). The concept of the structure matrix was extended by Finn (1976) to define a cycling index. Patten et al. (1976) produced further definitions of the structure matrix. Barber (1978) used the structure matrix in combination with Markovian analysis. Levine (1977) used the same matrix to outline an approach to niche overlap. Hannon (1976, 1979) and Herendeen (1981) introduced the concept of ecological cost, derived from the structure matrix, and Hannon (1976, 1979) developed an ecosystem optimization theory using this cost concept. Ulanowicz and Kemp (1979) used the feeding coefficient matrix to partition the various populations among the appropriate trophic levels. Constanza and Neill (1981) gave a linear programming solution to ecosystem production techniques using a variation of the structure matrix. Ulanowicz (1983) identified cycling structure. Hannon et al. (1983) argue that the ecological cost is the price or unit value in ecosystems and is the basis for ecological exchange. Finally, Hannon (1982) has blended this concept of price with that of age-weighted biomass to further extend the basis for ecological exchange.

In this paper I describe the most recent concepts for organizing ecosystems flow data and the methods for manipulating the resulting form to reveal the interdependence of the ecosystem processes. The format and analysis are useful for sensitivity and stability analysis, but perhaps their most important use is in the management of ecosystems. The procedure, although a linear, static approach, does allow an estimate of the total contribution of each process toward an extra unit of output of the product of any one of them. The analysis can be performed on any ecosystem, regardless of size, complexity of interconnection or number of inputs or outputs.

First, the flow accounting procedure is defined and described. Net input, production and net output are the three principle elements in the procedure. A theory is developed which produces a matrix of production flow interdependence and a set of weighting factors or unit values which distribute the net input over the net output. The procedure is derived in analogy to recent developments in economic theory. One of my long-term goals is to provide a method for comparing behavioral or optimality theories with experimental results. These theories could focus on growth maximizing strategies of the processes or on optimal feeding strategies, for example. The basic theory is applied to three marine ecosystems, and the results are discussed. A more useful approach to the same subject is developed in an appendix, but existing data sets are not detailed enough for application.

The Ecosystem Accounting Procedure

Before approaching the question of value measure in the ecosystem, one must lay out a flow accounting procedure to precisely define all flows and to show the unique origin and destination of these flows. Throughout this entire discussion all the flows are measured over a specific time period. The choice of the time period is crucial and must be determined mainly by biologists, based on their knowledge of the rate of change of the principal flows and the limitations upon measurement. A crucial factor in that decision is that a designated scarce net input (e.g. solar energy, water) remain scarce over the selected time period.

A schematic of the accounting scheme is shown in Fig. 1. The production matrix is $\boldsymbol{P}$,

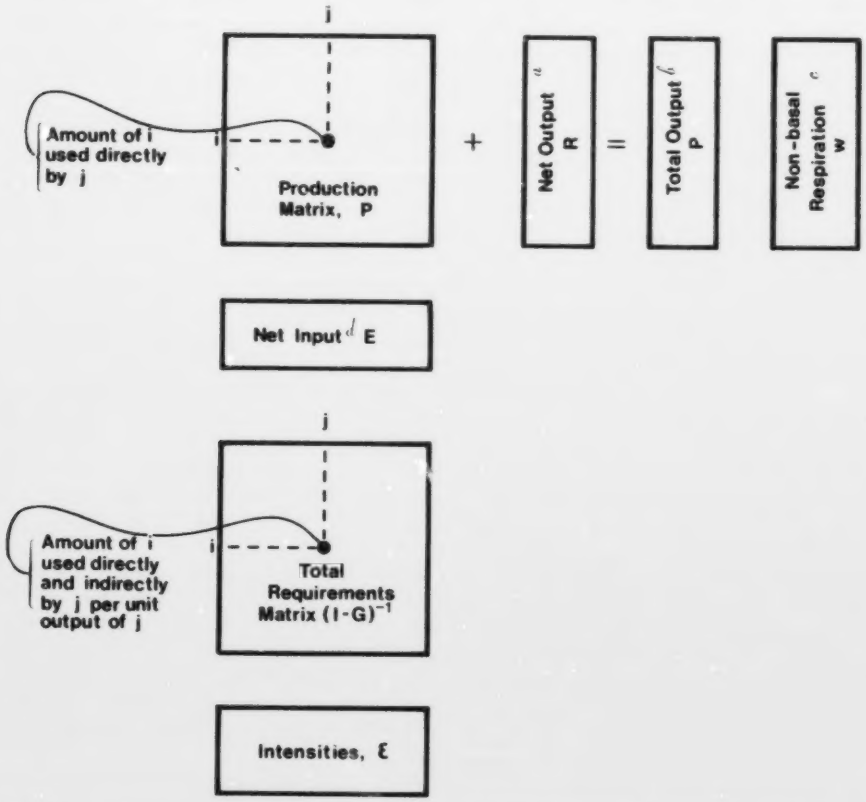


FIG. 1. The ecosystem accounting method. [a] The net output is composed of exports − imports + growth of stocks and inventory change + basal or resting metabolism (mass flow or free energy measure only); [b] The total output is the row sum of $\boldsymbol{P} + \boldsymbol{R}$; [c] The non-basal respiration rate represents the flow of heat and materials which occurs due to process activities such as pursuit, escape and reproduction. This vector is used only to demonstrate mass or energy balance was achieved in the system data (i.e., if all the flows were measured in the same units (and only each row must necessarily be so measured) then the column sum of $\boldsymbol{P} + \boldsymbol{E}$ is $\boldsymbol{P} + \boldsymbol{w}$); [d] The net input is composed of these flows which are not produced by the system, such as absorbed sunlight or initial endowments of nutrients.

each of whose elements is the amount of the row i commodity used by the column process j. At this point, each process (e.g., algae) is assumed to have only one unique output commodity (e.g., algae biomass), therefore the distinction between commodity and process is unnecessary. Later in this paper (Appendix A), however, I develop a slightly more complex accounting system to treat the more realistic case of multiple commodity output processes. The matrix $\boldsymbol{P}$ must be square to allow inversion.

The production exchange represented by $\boldsymbol{P}$, must have net inputs and net outputs. The net inputs are represented in a most general way as the matrix, $\boldsymbol{E}$, of *non-produced* commodities used by the system. For example, one of the row vectors of the net input matrix might be photosynthetically absorbed sunlight.[1] Another row might be the absorption of a special nutrient which was available as an initial endowment for the particular ecosystem. Each row of the net input matrix must have entries with the same units, but the units may vary from row to row. Some of the net inputs are more abundant than others. One of the net inputs is assumed to limit the system flows, and that input is designated as the vector $\boldsymbol{e}$. It might be the availability of light energy, water, nitrogen, or carbon, for example. Within a complex web of ecosystem exchanges it seems possible that several different net inputs might be simultaneously scarce. That is, while light may be limiting the production of the photosynthesizers, phosphorous may be limiting the growth of a carnivore, for example. This situation suggests that a hierarchy of scarce net inputs may exist — those limitations occurring nearest to the primary production process being the most important. Although only one net input vector will be designated as limiting the system growth, all the vectors of inputs will nonetheless be accounted for in the mass or energy balancing process to be described.

The net output is a series of vectors which describe the "surplus" production of each commodity beyond that level of output just needed to maintain the production processes themselves. The complete definition of the net output process depends to a certain extent on what one ultimately wants from the mathematical description. (See Appendix B.) The definition of net output used here is an attempt to provide a universally useful description — capable of accounting for all the important flows under a variety of system conditions such as growth, steady state or decline.

Accordingly, the net output matrix, $\boldsymbol{R}$, consists of a series of column vectors: (1) The gains or losses in biomass during the period, (2) The changes in the stocks of produced nonbiomass commodities (e.g., ammonia), (3) The amounts of the commodities exported less those imported across the system boundary[2], (4) The commodity flows which result from the "resting" metabolism of the processes. Sometimes these latter quantities are referred to collectively as the measure of basal metabolism, or the minimum respiration level. It is thought of here as the total respiration commodity vector less all respirations caused by predation, predation avoidance, and reproduction. This basal metabolism vector is intended to be a surrogate measure for the unavoidable stock decomposition occurring in the ecosystem due to the entropy forming nature of all processes. In this ecosystem description, basal metabolism is the vector of flows which indicates the amount of structural repair and rebuilding during the specified period. Such activity levels will depend on the organism's structural complexity and might be assumed to be proportional to the stock size. The basal metabolism flow should be measured in the same units as the

[1]The input need not be restricted to that which is *absorbed* photosynthetically. One might also include the energy absorbed for heating purposes, or even the reflected energy. For example, as succession progresses, the reflected quantities of energy are probably diminishing, and, for the purposes of comparison between stages, such quantities are relevant.

[2]Note that if an imported commodity is indistinguishable from a produced commodity, it is therefore not a *net* input; it is made a (negative) part of the net output.

(row) commodity of which it is a part. It should also represent the same energy quality as this row commodity (e.g., grams-carbon released due to basal respiration is also the flow measure of the produced biomass). The basal metabolism is the release of energy which was captured at an earlier time, and it theoretically is not a part of the net input. Part of the net input in the current period, of course, does go directly and indirectly toward building biomass to replace the decomposed biomass; and in the steady-state this part is equal exactly to the heat of basal metabolism. The remainder of the heat of respiration could be thought of as having zero free energy, therefore, its addition to $\boldsymbol{R}$ should change nothing. This heat may be considered as mechanical in origin (friction and the release of stored work) or as the heat released by those chemical reactions not specifically represented in the production and net output matrices. Therefore, these heat flows are not considered to be on the output side; but they are contained in the high quality flows on the input side of each process, as I show below.

The row sum of $\boldsymbol{R}$ is designated as $\boldsymbol{r}$. The remainders of the respiration outputs (seeking and avoiding prey and reproduction) are designated as the vector $\boldsymbol{w}$ (see Fig. 1).

One of the ways the system seems to cope with resource scarcity (or abundance) is to substitute among net inputs and among net outputs. Another way is to produce a given set of net outputs with less (or more) production and net input flows. This latter possibility, called "structural technological change" in economics, is very hard to model accurately. Short run technological change is the economist's term for succession. Long term technical change is evolution.

The total output vector $\boldsymbol{P}$ is *defined* as the row sum of the production and net output matrices. Row and column mass (or energy) balance is verified by summing down the column of a particular process (including all of the net inputs), and comparing the result to the sum across the row of the (single commodity) output of that process (including the net output and vector of predation and reproductive respiration). The column sums may not be calculable, however, if the commodities are measured in different units. The only requirement for the accounting system is that each commodity be represented across the row of $\boldsymbol{P} + \boldsymbol{R}$ in the same units, so that the total output vector can be formed. Note that although a column sum of $\boldsymbol{P}$ might be taken if all commodities were measured in the same units, this sum is not meaningful for the living ecosystem. The inputs to a given process do not necessarily have a value to the system in proportion to their mass or energy content. This point is made clear in equation (4) below.

The Theoretical Flow Relations

It is possible to derive a relationship between the net and total output flows provided certain assumptions are made about the unit flow ratios.

If the above accounting procedure has been adhered to, the following theoretical relationship obtains. Note that we have the definition,

$$(1) \qquad \boldsymbol{P}\boldsymbol{u} + \boldsymbol{r} = \boldsymbol{P} = \hat{\boldsymbol{P}}\boldsymbol{u},$$

where $\boldsymbol{u}$ is a vector of ones.
Solving for $\boldsymbol{r}$:

$$\begin{aligned}(2) \qquad \boldsymbol{r} &= (\hat{\boldsymbol{P}} - \boldsymbol{P})\boldsymbol{u} \\ &= (\hat{\boldsymbol{P}} - \boldsymbol{P})\hat{\boldsymbol{P}}^{-1}\boldsymbol{P} \\ &= (\boldsymbol{I} - \boldsymbol{P}\hat{\boldsymbol{P}}^{-1})\boldsymbol{P},\end{aligned}$$

where $\hat{}$ signifies a vector diagonalized to form a matrix. If $\boldsymbol{P}\hat{\boldsymbol{P}}^{-1} = \boldsymbol{G}$, we have a result similar to that derived by Hannon (1973), namely,

(3) $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{r}$,

except that in equation (3), $\boldsymbol{P}$ is not the *total* system output. Here $\boldsymbol{P}$ excludes the non-basal metabolism vector, $\boldsymbol{w}$. It is clear from equation (1) that if $\boldsymbol{w}$ were known, the $\boldsymbol{P}$ of equation (3) could be expanded to incorporate the total input requirements vector by simple vector addition ($\boldsymbol{P} + \boldsymbol{w}$). The utility of this latter form of equation (3) was explained in Hannon (1973).

If we can specify the net output changes, and if we can assume that $\boldsymbol{G}$ is constant, then the total output requirements for each component or process in the system can be calculated from equation (3). If we assume that associated increases in $\boldsymbol{w}$ are proportionate, then the increases are porportionate in total flow of each commodity can be estimated. The resulting commodity output requirements can be compared to the capabilities of each process to make the needed changes. For example, some processes may not be able to expand output levels beyond certain physical constraints, such as niche size (e.g., land area) or perhaps due to a shortage of a crucial nutrient (e.g., nitrogen). In some cases the constraint may be impractical to remove (e.g., land surface shortage), and in other cases external intervention may be possible (e.g., fertilizers). An example of this procedure is presented below.

So far the analyses have been made without specific reference to the external inputs. Yet the results of equation (3), coupled with information on process constraints, allows an analyst to estimate the system's ability to react to increased demand for its net output. The system constraints are synonymous with the ideas we hold about scarcity. A more useful approach to scarcity would be to slightly reformulate the treatment of net inputs. Such a recasting of the problem brings us very much closer to the economist's view of cost and value, and it opens up methods of testing various hypotheses of ecosystem behavior.

Since input scarcity is the controlling factor in ecosystem expansion, we should directly incorporate the system's net inputs into the analysis. Recall that the scarce net input vector is designated $\boldsymbol{e}$. Imagine a vector of commodity weighting factors, $\boldsymbol{\epsilon}$ (the intensities), which represents both the direct and indirect use of $\boldsymbol{e}$ in the production of a unit of each of the commodities. Then the equation which expresses the balance of use of $\boldsymbol{e}$ is,

(4) $\boldsymbol{\epsilon P} + \boldsymbol{e} = \boldsymbol{\epsilon}\hat{\boldsymbol{P}}$;

that is, the direct and embodied inputs of $\boldsymbol{e}$ equal the embodied outputs of $\boldsymbol{e}$, for each commodity. Note that while the total commodity output can be meaningfully measured, the column sum of $\boldsymbol{P}$ cannot be, unless the commodity entries are weighted with the unit values for the system ($\boldsymbol{\epsilon P}$). Equation (4) can be solved for these weights,

(5) $\boldsymbol{\epsilon} = \boldsymbol{e}\hat{\boldsymbol{P}}^{-1}[\boldsymbol{I} - \boldsymbol{G}]^{-1}$,

involving the same total requirements matrix as before. If $\boldsymbol{e}$ is actually the only scarce input, then the $\boldsymbol{\epsilon}$ is a vector of system values or prices.[3] They are as stable as the normalized terms, $\boldsymbol{e}\hat{\boldsymbol{P}}^{-1}$ and $\boldsymbol{G}$, are constant. These two terms are viewed in economic systems as describing the "technology" of production.[4]

---

[3]Providing that the measure of time cost of biomass is zero (see Appendix D).

[4]To economists, equation (4) or (5) describes an economic system where the total commodity cost ($\boldsymbol{\epsilon P} + \boldsymbol{e}$) is directly proportional to output $\boldsymbol{P}$. Therefore, average and marginal commodity costs are equal (and equal to price, $\boldsymbol{\epsilon}$). Marginal cost is the unit cost of producing one more unit; average cost is the total cost divided by all the units produced. Normally the marginal cost is lower, but in the steady state these two costs are equal. This means that the competitive, profit-maximizing economy is at long-term equilibrium (economic profits are zero for every producer), and therefore the price ($\boldsymbol{\epsilon}$) is detectable from the production process alone. Only the level of output is set by the consumers. This simple and convenient view of the system should serve mathematical ecology as a suitable entry point for the (hopefully) useful views of theoretical economics. Only when the concept of price and value is correctly interpreted in ecology will the older science begin to be of value (Hannon et al. 1983).

The ϵ_i calculated in this way also show how much more of the net input must be supplied to produce one more unit of output, P_i (or r_j). For example, if solar energy is the only net input to the (steady state) ecosystem, then ϵ_{fish} is the amount of solar energy needed directly and indirectly to allow the production for export of one more unit of fish. Since the solar energy absorption rate is also equal to the waste heat exhaust rate, the ϵ are tied to the definition of waste heat (see Appendix B).

If the medium of flow represented in matrix $\boldsymbol{P}$ (such as carbon or energy) is known to be associated with certain trace elements or compounds, then one of the input vectors of $\boldsymbol{E}$ could represent the amounts of that tracer and its points of introduction. Accordingly, with no tracer losses, the resulting $\boldsymbol{\epsilon}$ vector should represent the concentrations of that tracer in the various commodities.

It is shown in Appendix B that if the non-basal respiration $\boldsymbol{w}$ is made a formal part of the net output, then $\boldsymbol{\epsilon} = \boldsymbol{u}$, the vector of ones. On the other hand, if none of the respiration (basal and nonbasal) is included in the net output, and if the system has no net export and no growth of any commodity, then the intensities become infinite. The inclusion of the basal metabolism as a part of the net output is appropriate, as noted above, but it is now clear why this is also mathematically necessary.

The final step in the development of this theory is to demonstrate that the intensities distribute the scarce net input exactly over the net output. Rewriting equation (4) gives

$$\boldsymbol{\epsilon}(\hat{\boldsymbol{P}} - \boldsymbol{P}) = \boldsymbol{e}, \tag{6}$$

$$\boldsymbol{\epsilon}(\hat{\boldsymbol{P}} - \boldsymbol{P})\boldsymbol{u} = \boldsymbol{\epsilon}\boldsymbol{u} = \boldsymbol{e}^T = \text{total input of } \boldsymbol{e};$$

$$\boldsymbol{\epsilon} \cdot \boldsymbol{r} = \boldsymbol{e}^T, \text{ since } (\hat{\boldsymbol{P}} - \boldsymbol{P})\boldsymbol{u} = \boldsymbol{r}.$$

Thus, the total net input of $\boldsymbol{e}$ into the ecosystem is the net output weighted by the intensities.

The possible utility of the theory in ecosystem management and analysis is shown in the examples below. The basic problem arises that in reality most processes have joint or multiple commodity outputs. In equations (1) through (6) the processes were defined as having only one output. This is the way most data collections are arranged (although somewhat artificially so). This problem can be avoided by defining two matrices: a "Use" matrix similar to the $\boldsymbol{P}$ matrix and a "Make" matrix indicating which processes made how much of which commodity. The revised theory, similar to the development above, is presented in Appendix A. Although the available data are arranged as single commodity products of their processes, the improved theory should be applied to future experimental research.

The theory introduced in this section and in Appendix A has two general applications, which will not be pursued further here due to lack of data. It is used in the development of the maximum energy storage hypothesis and the optimum feeding choice theory as outlined in Hannon (1979). The present treatment of intensity theory is more meaningful than the one given in the 1979 paper, but these two previous applications of the theory are still worthy goals. Basically, I am trying to predict the choices made by the components of a living system. The energy storage maximizing theory and the optimal feeding strategy are just two of many guesses at what gives rise to that behavior. As these guesses are based on the concept of the direct and indirect intensity, they are unlike any other ecosystem theory of which I am aware.

Application of the Theory to Available Marine Data Sets

Since the first application of input−output analysis to ecological data, a surprisingly large number of data sets on ecosystem flows have emerged. I have chosen three of these

sets to exemplify how sets of flow measures may vary in type and complexity. Unfortunately, reported net outputs usually do not specify the basal metabolism levels. As a result, in all the following examples the net outputs are composed of exports from the system and the variation of the ϵ over u is due to the *total* respiration. Since some of this respiration is non-basal, variation in the calculated ϵ is quite large.

The first of the example systems is shown in Fig. 2, which represents the flows of energy in the North Sea Marine ecosystem (Steele 1974). Several approximations had to be made: the amount of absorbed solar energy was estimated; the herbivore fecal output had to be arbitrarily divided between the macro and micro benthos; and the respiration of the large fish was approximated. Otherwise, data for the 10 component system were estimated by Steele from available measurements. The system absorbs about 9000 kcal $m^{-2} \cdot yr^{-1}$ and annually exports about 6 kcal $\cdot m^{-2} \cdot yr^{-1}$ as protein.

The production matrix is shown in Table 1a, and the resulting $(I - G)^{-1}$, or "Total Requirements" matrix, is given in Table 1b. The energy intensities, or energy flow weighting factors, ϵ, are shown at the bottom of Table 1b. They vary by a factor of almost 3000; from 10 for the algae to 28360 for the large fish. This system can be thought of as two systems — the first is comprised of processes 1, 2, 3 and 4, which contribute two export flows to the rest of the processes comprising the second system. This division helps to explain how a large variation in the energy intensities between pelagic and large fish could exist. The ϵ represent the ecosystem prices or unit values if sunlight were the constraining net input. If sunlight is actually controlling, then the system commodities are valued in the following increasing order: 1, 2, 5, 7, 6, 3, 4, 9, 8, 10, with algae the lowest and the large fish the highest. It would be very interesting to compare these values with

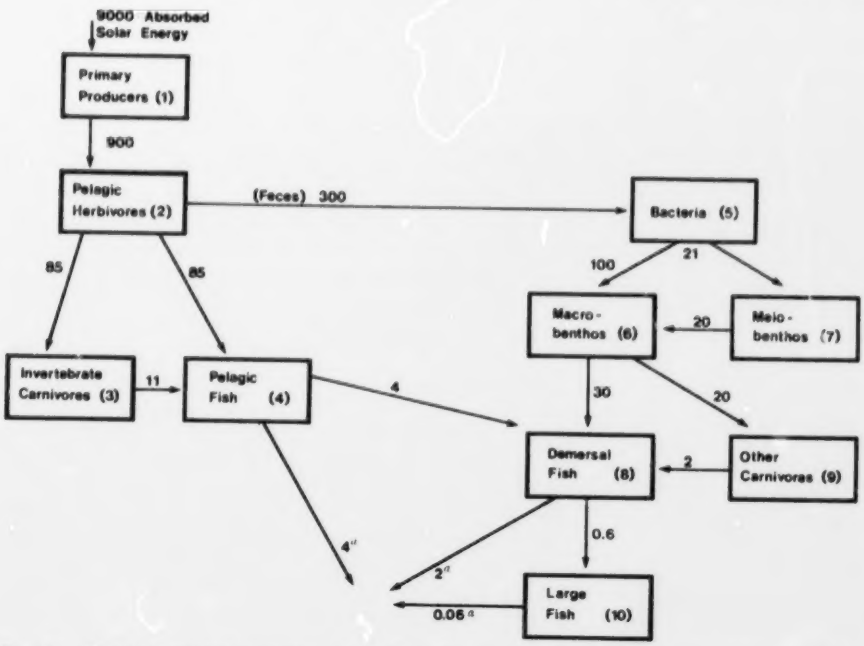


FIG. 2. North Sea marine ecosystem energy flow diagram (Steele 1974). Units are kcal $\cdot m^{-2} \cdot yr^{-1}$. (*Exported to fishermen.)

TABLE 1. The production (1a) and total requirements (1b) matrices for a marine ecosystem (Steele 1974). In Table 1a, Nos. in parenthesis are $\hat{e} \cdot P$. The total energy input is their sum. Units are kcal·m^{-2}·yr^{-1}. e are dimensionless.

Table 1a

	1	2	3	4	5	6	7	8	9	10	Export	Total output	Respiration
1 Algae	—	900 (9000)	—	—	—	—	—	—	—	—	—	900	8100
2 Herb.	—	—	85 (1628)	85 (1628)	300 (5745)	—	—	—	—	—	—	470	430
3 Invert. carn.	—	—	—	11 (1650)	—	—	—	—	—	—	—	11	74
4 Pelag. fish	—	—	—	—	—	—	—	4 (1630)	—	—	4 (1630)	8	88
5 Bacteria	—	—	—	—	—	100 (4747)	21 (997)	—	—	—	—	121	179
6 Macrobenthos	—	—	—	—	—	—	—	30 (3446)	20 (2298)	—	—	50	70
7 Meiobenthos	—	—	—	—	—	20 (997)	—	—	—	—	—	20	1
8 Dem. fish	—	—	—	—	—	—	—	—	—	0.6 (1702)	2 (5670)	2.6	33.4
9 Other carn.	—	—	—	—	—	—	—	2 (2298)	—	—	—	2	18
10 Large fish	—	—	—	—	—	—	—	—	—	—	0.06 (1700)	0.06	0.54
E; Solar input	9000	—	—	—	—	—	—	—	—	—			
Total energy input	9000	9000	1628	3278	5745	5744	997	7374	2298	1702	6.06 (9000)		

Table 1b

	1	2	3	4	5	6	7	8	9	10
1	1.0	1.9150	14.798	40.752	4.747	11.488	4.985	283.63	114.88	2836.3
2	—	1.0	7.727	21.281	2.479	5.999	2.603	148.11	59.992	1481.1
3	—	—	1.0	1.379	—	—	—	2.122	—	21.22
4	—	—	—	1.0	—	—	—	1.539	—	15.39
5	—	—	—	—	1.0	2.42	1.05	46.53	24.20	465.34
6	—	—	—	—	—	1.0	—	19.23	10.00	192.29
7	—	—	—	—	—	0.4	1.0	7.692	4.0	76.92
8	—	—	—	—	—	—	—	1.0	—	10.00
9	—	—	—	—	—	—	—	0.769	1.0	7.69
10	—	—	—	—	—	—	—	—	—	1.0
Energy intensity, e	10.0	19.15	147.97	407.52	47.47	114.88	49.85	2836.3	1148.9	28363.0

the free energies of the various commodities. It is possible that the components of ϵ represent useful substitutes for the free energies of the commodities by virtue of symbolizing the actual energy needed directly and indirectly to produce them.

The terms in the dot product $\epsilon \cdot r$ indicate the amount of absorbed solar energy being absorbed and exported. These values are shown in parentheses in the "exports" column of Table 1a. The measured export of the demersal fish, although only half as large as the pelagic fish export, actually represents a solar energy export almost twice as large as the corresponding solar export from the pelagic fish. The relatively small (0.06 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}$) export of large fish represents an absorbed solar energy larger than that of the pelagic fish. Equation (6) is verified by the sum of the entries in parenthesis in the export column (9000), which is equal to the energy input from the sun.

The entries in any column of the total requirements matrix (Table 1b) indicate the direct and indirect production energy flow through each row commodity. For example, a unit export of demersal fish required a 148 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}$ flow through the herbivores, even though no direct flow occurred between these two processes. There are six separate pathways connecting these two processes, and equation (5) sums the appropriate combination of each into a single number representing the interdependency. As a further example, suppose one wished to export one more unit of "large fish." From Table 1b, column 10, we find that this increase would require 2836 more units of production flow through the algae, 1481 more through the herbivores, and so on — exclusive of needed respiration increases. The value of ϵ_{10} (28363) times 1.0 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}$ is the total increased need in respiration flow (and solar input). The individual increases in respiration (rounded off: 25530, 1350, 142, 168, 686, 269, 4, 128, 69, 9) required for one more unit of exported "large fish" can be found by multiplying ratios of the needed increases in total output (Table 1b, column 10) to the total output of each commodity, times the commodity respiration vector (last column, Table 1a). The sum of these respiratory increases is 28363, which is ϵ_{10}.[5] Therefore, the sum of the increases in *total input* to each process is the sum of the two vectors, or (rounded off: 28400, 2840, 164, 185, 1150, 462, 81, 139, 77, 10). This new vector is based on the assumption that the respiration and the total output are linearly related. This linearity is consistent with the assumption of constant ϵ. If the change in respiration can be obtained from an allometric knowledge of the specific processes, that procedure may be more desirable (Peters 1983). The general method for determining the change in total input to each process is given in Appendix C. Note that in Appendix C, the elements of ϵ are shown to be the total net input to the primary producer (the first column of the table) required for the respective processes to produce one more unit of net output. For example, $\epsilon_8 = 2840$ $\text{kcal} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$, which is the increase in net input to the primary producer required for the Pelagic Fish (8) to produce one more unit of export to the fishermen.

The ecologist must verify whether the system has the capacity for these increases, or whether it is possible for outside intervention to release natural constraints in the system to allow these flows — before the additional unit of export can be realized. Hence, what is needed is an understanding of the functional relationship between each commodity stock and its respiration or total input. Perhaps there is a limit keeping a certain process stock from achieving the desired respiration level. That process limit would then control the level of the possible increase in "large fish" export.

A complication arises if the basal metabolism (r_m) were included in the net output. The stocks would then need to increase for two reasons, first to handle the production and non-basal metabolism increases outlined above and a second time to accommodate $\epsilon \cdot r_m$.

[5] Actually $\epsilon_{10}{}^{-1}$; see Appendix C, equation (C1).

the non-basal metabolism increase needed for the increased basal metabolism. This process is iterative and reaches a limit if the basal fraction of total respiration shrinks as respiration increases. The iterative process is described in Appendix C.

The numbers in parenthesis in the production matrix (Table 1a) are $\hat{\epsilon} \cdot \boldsymbol{P}$, representing the direct plus the indirect energy embodied in the actual direct flow between a column process and row commodity input. For example, demersal fish receive 2298 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}$ of solar energy embodied in the 2 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}$ of direct flow from "other carnivores."

This matrix gives the relative importance of the direct to the indirect connections in the production matrix. As such, these results are useful in sensitivity analyses and studies of system stability or structural change with regard, for example, to increasingly scarce net input. The column sums of this matrix are shown at the bottom of Table 1a.

If the data for the production flow matrix $\boldsymbol{P}$ are known within certain limits of accuracy, the impact of variations in $\boldsymbol{P}$ on the production inputs $\boldsymbol{P}$, the total input vector, (see Appendix C) and ϵ can be calculated rather easily using Monte Carlo techniques. From such an analysis one could guide research toward improving experimental accuracy in those parameters where greater precision would be most efficacious. For example, if greater accuracy in the total production flow P_1 were desired, the Monte Carlo technique would show those variations in $\boldsymbol{P}$ which contributed most to variations in P_1. On the other hand, the research team may find that expense of reducing the most sensitive uncertainties in $\boldsymbol{P}$ would be very great. If the dollar cost of reducing the uncertainty in each element of $\boldsymbol{P}$ by one unit were known, then it would be possible to construct a cost-minimizing research strategy which would improve the accuracy of P_1 by a desired amount.

The intertidal oyster reef ecosystem (Dame and Patten 1981) is shown in Fig. 3. This is a more complex system than the one depicted in the Steele diagram. The filter feeders (oysters) receive the system's sole source of energy in the form of phytoplankton and suspended organic matter. The system also contains a detritus component which in the input−output perspective is indistinguishable from a living component: it takes in and gives off organic materials. Mortality and resuspended material are considered as part of the net system output by the authors, and I follow that criterion.

The data are arranged in the appropriate manner in Table 2a. The resulting total requirements matrix (Table 2b) is much fewer null entries than in the Steele system, primarily because of the interconnections of processes 2 and 4, and 2 and 5. The ϵ (Table 2b) vary by more than a factor of 9. The predators do not have the highest intensity value, mainly because they feed principally on the oysters (herbivores). The principal feeder on the detritus, the microbiota, have the highest ϵ value. Export of one unit of predator would cause the production output increases given in column 6 of the $(\boldsymbol{I} - \boldsymbol{G})^{-1}$ matrix in Table 2b and a total respiration increase of 7.50 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ (ϵ_6). These flow increases include a 1.072 $\text{kcal} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ increase in the predator cycling. This means that to produce one more unit of predator requires an additional 7.2% increase in indirect predator self consumption. In fact, all processes 2 through 6 have diagonal terms (indirect self consumption) which are larger than one, even though no direct self consumption occurs in any process. This indirect self consumption is attributable to the high degree of feedback easily recognized from Fig. 3.

The Crystal River Tidal Marsh (Homer and Kemp 1983) is an example of one of the more complex ecosystems ever described in input−output terms. The system is shown schematically in Fig. 4. This system also has a detritus sector, which I treat as an important process in the ecosystem. It contains micro-organisms which break down certain compounds for recycling. If this sector were placed in the net output column, then its outputs would have to appear in the net input rows. If these inputs were not limiting system growth

TABLE 2. The production (2a) and total requirements (2b) matrices for an intertidal oyster reef (Dame Patton 1981). Units are kcal $m^{-2} \cdot day^{-1}$. ϵ are dimensionless.

Table 2a

	1	2	3	4	5	6	Exports resuspen. mortality	Total output[a]	Respiration
Oysters (1)	0	15.79	0	0	0	0.51	10.44	26.74	14.73
Detritus (2)	0	0	8.17	7.27	0.64	0	6.19 (resuspen.)	22.27	—
Microbiota (3)	0	0	0	1.21	1.21	0	0	2.42	5.75
Meiofauna (4)	0	4.24	0	0	0.66	0	0	4.90	3.58
Deposit feeders (5)	0	1.91	0	0	0	0.17	0	2.08	0.43
Predators (6)	0	0.33	0	0	0	0	0.05	0.38	0.30
Phytoplankton and susp. organ[b]	41.47	0	0	0	0	0			

Table 2b

	1	2	3	4	5	6
1	1.0	2.597	8.767	6.019	7.810	4.833
2	0	3.561	12.02	8.255	10.71	4.788
3	0	0.385	2.300	1.140	1.819	0.813
4	0	0.781	2.638	2.811	2.667	1.192
5	0	0.330	1.115	0.765	1.993	0.891
6	0	0.053	0.180	0.124	0.161	1.072
Energy intensity, ϵ	1.55	4.03	13.6	9.34	12.1	7.50

[a]Total output is the flow matrix row sum plus net output (exports).
[b]Biomass absorbed by the oysters. Required solar input is unknown.

in some way, their presence would not be felt in the system model. This would cause a great loss in the model detail and meaning. If these inputs were limiting, the solar input would be neglected. The detritus sector belongs in the matrix as though it were a living process. It is virtually the only point of feedback in this system. (There is a minor cycle between Pinfish and Needlefish.)

The important aspect of this system is that data have been collected on a nearby, practically identical tidal marsh which is subjected an elevated temperature (+6°C) caused by a power plant effluent.

The production matrices for the normal and disturbed marsh systems are given in Table 3, and the total requirements matrices in Table 4. At the lower edge of Table 4, the energy intensities and their ratios are given. In every process the intensities rose when the system was disturbed by heating. The net output dropped 33% from 1265 to 854 $kcal \cdot m^{-2} \cdot yr^{-1}$. The intensity for Stingray (6), for example, increased by 11-fold, while the microphyte intensity increased by only 2%, indicating the relative severity of the heating on the Stingray.

The heated tidal marsh is obviously less productive. For example, the macro- and microphytes absorbed 18% less energy in the heated system. What is not obvious is that the energy intensities of every process rose as well, particularly among those processes which did not absorb solar energy directly. Although those processes which are separated the farthest from the plants reduced their direct dependence on primary producers, they nonetheless *increased* their total (direct and indirect) dependence on these autotrophs. The

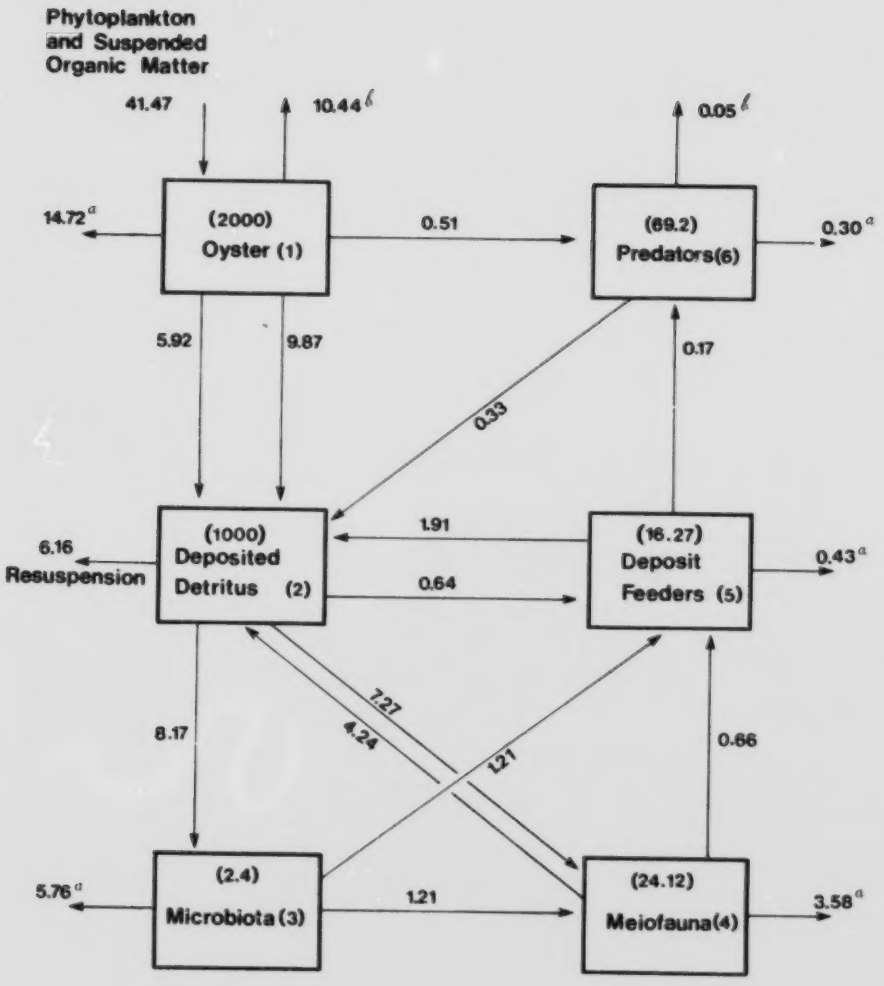


FIG. 3. Intertidal oyster reef ecosystem energy flow diagram (Dame and Patton 1981). Flow units are kcal · m^{-2} · day^{-1}, stock units kcal/m^2. ([a] respiration; [b] mortality flows leaving the system.)

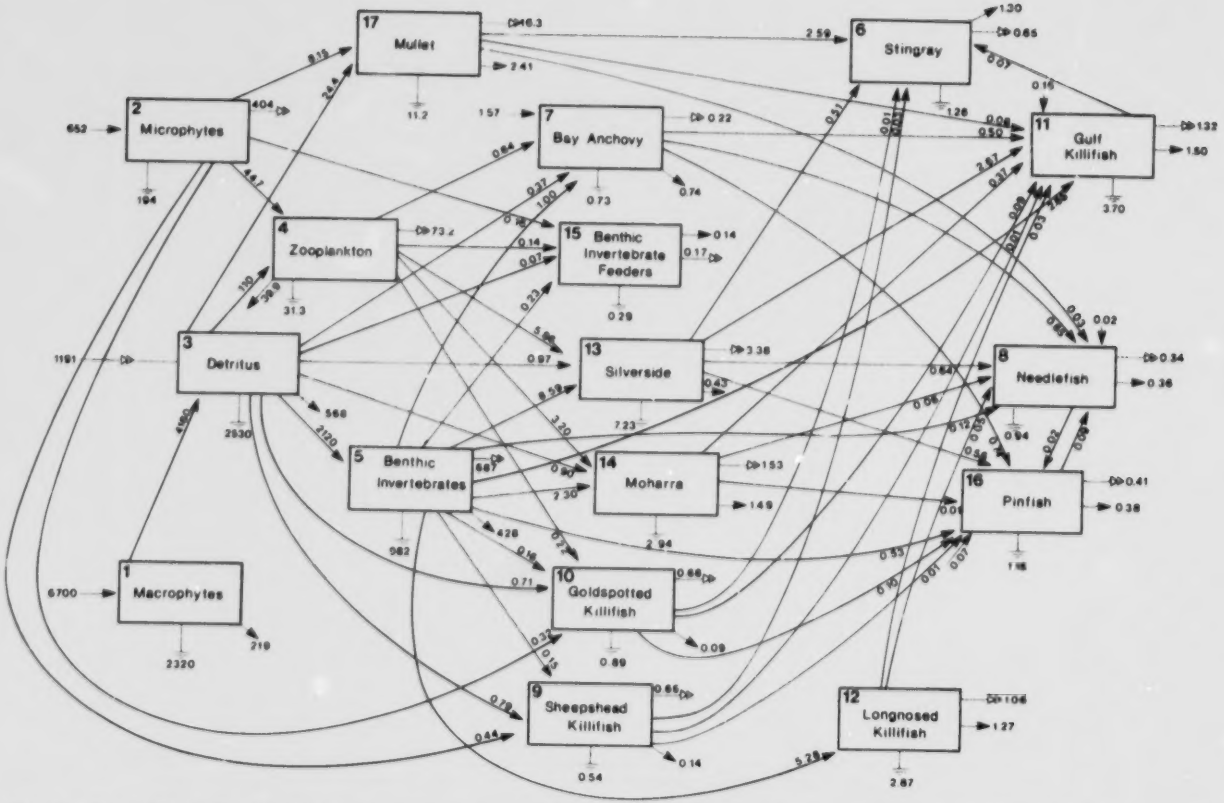


FIG. 4. Crystal River (Florida) tidal marsh ecosystem carbon flow diagram (Homer and Kemp 1983). Units are mg C · m $^{-2}$ · day $^{-1}$. Ground symbol is respiration. (By permission from *Math. Biosci.*, Vol. 64, p. 231 (1983) by Elsevier Science Publishing Co., Inc.)

TABLE 3. The production matrix for the undisturbed (disturbed, +6°C) tidal marsh ecosystem Crystal River, Florida (Homer and Kemp 1983). Units are g C·m $^{-2}$·day $^{-1}$, except as noted.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Exports	Total output	Respiration
Microphytes	1	0	0	404	44.7	0	0	0	0	.44	.32	0	0	0	0	.16	0	8.15	0	457.77	194
		(0)	(0)	(319)	(39.05)	(0)	(0)	(0)	(0)	(.12)	(.10)	(0)	(0)	(0)	(0)	(0)	(0)	(6.08)	(.65)	(365)	(165)
Macrophytes	2	0	0	4163	0	0	0	0	0	0	0	0	0	0	0	0	0	0	219	4382.	2322
		(0)	(0)	(3156)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(166)	(3322)	(2166)
Detritus	3	0	0	0	109.8	2117.9	0	.37	0	.79	.71	0	0	.97	.90	.07	0	24.4	567.79	2823.7	2530.1
		(0)	(0)	(0)	(117.15)	(1485.1)	0	(7.74)	(0)	(.22)	(.22)	(0)	(0)	(.06)	(.11)	(0)	(0)	(17.19)	(616.51)	(2244.3)	(2070.3)
Zooplankton	4	0	0	73.2	0	0	0	.64	0	0	.22	0	0	5.86	3.2	.14	0	0	39.94	123.2	31.30
		(0)	(0)	(78.1)	(0)	(0)	(0)	(13.6)	(0)	(0)	(.07)	(0)	(0)	(.33)	(.4)	(0)	(0)	(0)	(32.4)	(124.9)	(31.30)
Benthic invertebrates	5	0	0	686.9	0	0	0	1.0	.12	.15	.61	2.68	5.28	8.59	2.3	2.3	.53	0	427.91	1136.3	981.6
		(0)	(0)	(742.6)	(0)	(0)	(0)	(2.1)	(0)	(.04)	(.19)	(1.71)	(3.48)	(.49)	(.28)	(2.38)	(.09)	(0)	(33.55)	(786.91)	(698.19)
Stingray	6	0	0	.65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.30	1.95	1.26
		(0)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.01)	(.04)
Bay Anchovy	7	0	0	.22	0	0	0	0	.65	0	0	.50	0	0	0	0	.74	0	.74	2.85	.73
		(0)	(0)	(4.69)	(0)	(0)	(.04)	(0)	(.24)	(0)	(0)	(.37)	(0)	(0)	(0)	(0)	(.23)	(0)	(.75)	(6.32)	(17.12)
Needlefish	8	0	0	.34	0	0	0	0	0	0	0	0	0	0	0	0	.02	0	.36	0.72	94
		(0)	(0)	(.06)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.01)	(.07)	(.23)
Sheepshead killifish	9	0	0	.65	0	0	.03	0	0	0	0	.01	0	0	0	0	.01	0	.14	0.84	.54
		(0)	(0)	(.18)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.19)	(.19)
Gold spotted killifish	10	0	0	.68	0	0	.01	0	0	0	0	.09	0	0	0	0	.10	0	.09	0.97	.89
		(0)	(0)	(.21)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(.1)	(.32)	(.26)
Gulf killifish	11	0	0	1.32	0	0	.07	0	0	0	0	0	0	0	0	0	0	0	1.50	2.89	3.70
		(0)	(0)	(.54)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(.73)	1.28	(1.39)
Longnosed killifish	12	0	0	1.06	0	0	0	0	.05	0	0	.03	0	0	0	0	0	0	1.27	2.41	2.87
		(0)	(0)	(.70)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.17)	(0)	(0)	(0)	(0)	(0)	(0)	(.69)	(1.56)	(1.92)
Silverside	13	0	0	3.38	0	0	.51	0	.64	0	0	2.67	0	0	0	0	.56	0	.43	8.19	7.23
		(0)	(0)	(.18)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.33)	(0)	(0)	(0)	(0)	(0)	(0)	(.05)	(.56)	(.32)
Mohorra	14	0	0	1.53	0	0	0	0	.06	0	0	.37	0	0	0	0	.01	0	1.49	3.46	2.94
		(0)	(0)	(.16)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.25)	(.41)	(.38)
Benthic invert. feeders	15	0	0	.17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.14	0.31	.29
		(0)	(0)	(.47)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.52)	(.99)	(1.39)
Pinfish	16	0	0	.41	0	0	0	0	.09	0	0	0	0	0	0	0	0	0	.38	0.88	1.16
		(0)	(0)	(.07)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.09)	(.16)	(.18)
Mullet	17	0	0	16.29	0	0	2.59	0	.03	0	0	.08	0	0	0	0	0	0	2.41	21.4	11.15
		(0)	(0)	(11.64)	(0)	(0)	(.01)	(0)	(.06)	(0)	(0)	(.06)	(0)	(0)	(0)	(0)	(0)	(0)	(1.37)	(13.14)	(10.13)
Absorbed energy input*		6517.7 (5300)	67040. (54880)	0 (0)	0 (0)	0 (0)	0 (0)	15.7 (0)	0.2 (0)	0 (0)	0 (0)	1.6 (0)	0 (0)	0 (0)	0 (0)	0 (0)	.7 (.2)	0 (0)	1265 854		

*kcal·m $^{-2}$·day $^{-1}$. Carbon flows converted at 10 kcal/g C. Original data are in mass flow balance.

TABLE 4. The total requirements matrix for the undisturbed (disturbed, +6°C) tidal marsh ecosystem Crystal and River, Florida (Homer and Kemp 1983). The energy intensities, ε, and the ratio of the intensities for the disturbed undisturbed ecosystems; dimensionless except εP^{-1} and ε are kcal/g C.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Microphytes	1	1	0	.30	.63	.56	1.33	.38	1.78	.91	1.05	1.78	1.23	1.08	1.04	1.29	1.52	.72
		(1)	(0)	(.49)	(.77)	(.92)	(11.3)	(2.55)	(9.68)	(1.39)	(1.36)	(2.66)	(2.04)	(1.31)	(1.50)	(2.20)	(4.18)	(1.10)
Macrophytes	2	0	1	2.87	2.55	5.34	6.96	2.82	13.30	3.65	6.04	13.88	11.70	7.77	6.66	5.77	11.64	3.27
		(0)	(1)	(4.34)	(4.07)	(8.2)	(72.9)	(16.8)	(62.5)	(6.75)	(8.74)	(21.4)	(18.3)	(10.0)	(10.7)	(19.7)	(28.77)	(5.68)
Detritus	3	0	0	1.94	1.73	3.62	4.72	1.91	9.02	2.48	4.10	9.42	7.94	5.27	4.52	3.91	7.90	2.22
		(0)	(0)	(3.09)	(2.90)	(5.83)	(51.8)	(11.95)	(44.4)	(4.80)	(6.22)	(15.2)	(13.0)	(7.14)	(7.64)	(14.0)	(20.5)	(4.04)
Zooplankton	4	0	0	.05	1.05	.10	.34	.28	1.25	.07	.34	1.08	.21	.86	1.06	.56	.91	.06
		(0)	(0)	(.12)	(1.12)	(.23)	(10.7)	(2.63)	(9.14)	(.19)	(.47)	(1.38)	(.52)	(.87)	(1.28)	(.56)	(3.91)	(.16)
Benthic invertebrates	5	0	0	.47	.42	1.88	1.51	.82	4.04	.78	1.63	4.38	4.13	2.34	1.77	1.70	3.61	.54
		(0)	(0)	(1.03)	(.97)	(2.94)	(18.6)	(4.32)	(16.0)	(1.81)	(2.67)	(7.04)	(6.57)	(3.25)	(3.23)	(7.08)	(7.86)	(1.35)
Stingray	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
		(0)	(0)	(0)	(0)	(0)	(1)	(0)	(0)	(1)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Bay anchovy	7	0	0	0	0	0	.01	1	1.01	0	0	.17	0	0	0	0	.86	0
		(0)	(0)	(.01)	(.01)	(.01)	(4.11)	(1.03)	(3.53)	(.01)	(.01)	(.32)	(.03)	(.02)	(.02)	(.03)	(1.48)	(8.72)
Needlefish	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	.02	0
		(0)	(0)	(0)	(0)	(0)	(0)	(0)	(1)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Sheepshead killifish	9	0	0	0	0	0	.02	0	0	1	0	0	0	0	0	0	.01	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(.01)	(1)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)
Gold Spotted killifish	10	0	0	0	0	0	.01	0	.01	0	1	.03	0	0	0	0	.11	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(0)	(0)	(1)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)
Gulf killifish	11	0	0	0	0	0	.04	0	0	0	0	1	0	0	0	0	0	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(.01)	(0)	(0)	(1.01)	(0)	(0)	(0)	(0)	(0)	(0)
Longnosed killifish	12	0	0	0	0	0	0	0	.07	0	0	.01	1	0	0	0	0	0
		(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.02)	(0)	(0)	(.14)	(1)	(0)	(0)	(.01)	(.01)	(0)
Silverside	13	0	0	0	0	0	.30	0	.98	0	0	.93	0	1.01	0	0	.67	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(.01)	(0)	(0)	(.26)	(0)	(1)	(0)	(0)	(0)	(0)
Mojarra	14	0	0	0	0	0	0	0	.08	0	0	.13	0	0	1	0	1.29	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(1)	(0)	(0)	(0)
Benthic Inv. feeders	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
		(0)	(0)	(0)	(0)	(0)	(.01)	(0)	(.01)	(0)	(0)	(0)	(0)	(0)	(0)	(1)	(0)	(0)
Pinfish	16	0	0	0	0	0	0	0	.13	0	0	0	0	0	0	0	1	0
		(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(1)	(0)
Mullett	17	0	0	.01	.01	.02	1.36	.01	.10	.01	.02	.08	.05	.03	.03	.02	.05	1.01
		(3)	(0)	(.02)	(.02)	(.03)	(1.27)	(.06)	(1.09)	(.03)	(.03)	(.13)	(.07)	(.04)	(.04)	(.01)	(.11)	(1.02)
εP^{-1}		14.24	15.30	0	0	0	0	5.51	42	0	0	55	0	0	0	0	80	0
		(14.52)	(16.52)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(.13)	(0)
ε		14.24	15.30	48.13	48.04	89.73	125.40	54.06	235.0	68.77	107.3	239.2	196.6	134.2	116.6	106.5	205.3	60.30
		(14.52)	(16.52)	(78.78)	(78.42)	(14.83)	(1368)	(314.6)	(1047)	(131.7)	(164.1)	(391.8)	(331.7)	(180.3)	(199.2)	(357.4)	(537.2)	(108.4)
ε ratio Dist./Undist.		1.02	1.08	1.64	1.63	1.66	10.91	5.82	4.46	1.92	1.53	1.64	1.69	1.34	1.71	3.36	2.62	1.80

process most severely damaged by heating is the Stingray (6). It fed mainly on Mullet (17) and Silverside (13), and heating reduced these direct inputs almost to zero (Table 3). But in Table 4, we see that although indirect Stingray dependence on Silverside was reduced almost to zero, the total dependence on Mullet was diminished by only 7%. This contrast occurred because the Mullet feeds on microphytes and detritus, and consequently was able to survive the heating impact; whereas the Silverside is farther removed from the basic food sources and almost ceased to exist in the warmer marsh gut (total output dropped from 8.19 to 0.56). From this examination I conclude that the Silverside was much more important to the Stingray than was the Mullet, even though its direct consumption of Mullet was 5 times greater than its consumption of Silverside. This conclusion is supported by the fact that the energy intensity of the Silverside in the undisturbed system is more than twice as high as the intensity of the Mullet, indicating that Silverside biomass probably represents a greater complex of nutrients and calories than does the Mullet.[6]

Ulanowicz (1984) points out that the Bay Anchovy was the most successful in increasing its biomass as a result of the disturbance. Its high ratio of ϵ values (disturbed to undisturbed) indicates that it did so by substantially increasing its dependence on the rest of the system and on the external or net inputs.

In general, the extent of the heating was so great as to change the structure, $(I - G)^{-1}$, of the ecosystem. Structural change is also indicated by the change in the energy intensities. This structural change makes it impossible to estimate the change in production and respiration flows due to a net output change (as was done in the North Sea ecosystem above). Also, the heating affected nearly all of the processes simultaneously, and so the net effect on each process was to some extent additive. The total effect of this much heating drove the system beyond the limits of simple linear analysis: using the methods described in this paper, I could not have predicted the heated condition from the unheated one — the change was too great.

The temperature level of the environment, like the available sunlight, is a major control variable in ecosystems. In fact, these are the only control variables in a thermodynamically closed ecosystem. A speculative connection is drawn in Appendix D between the entropy of the respiratory flows and the historic value of the relevant absorbed sunlight. It is because of this connection that the stocks of the biomass are related to the dynamic outputs and values.

Summary and Conclusions

An accounting procedure for the flows in a general ecosystem has been developed. The procedure is static in the sense that it applies to the flows during a specified time period. There are three basic flow sets: net inputs, production, and net outputs. From this procedure it is possible to determine the increases in production and in net (non-produced) input required for a desired incremental increase in net output. Only scarce net inputs are of interest. A scarce input is one which, if the level of its availability is changed, the flows in the system will change. Only one type of input is considered scarce in each period.

A simplified form of the theory is applied to three data sets on marine ecosystems. The three increasingly complex systems were used to demonstrate: (1) the nature of feedback loops in ecosystems; (2) the unexpected ranking of the importance of certain commodity flows in the ecosystem; (3) the method for estimating system-wide flow increases associated with small exports of a particular commodity; and (4) the nature of the system disturbance (caused by heating the system environment).

---

[6]One must also note that Sharks are present in the undisturbed system and they are grouped with the Stingray. The Sharks disappeared with heating. Therefore, the emphasis on the Silverside may be misplaced.

I found that the direct energy or mass flow was probably not the most appropriate measure of importance of a given exchange to the system. The direct plus indirect measure of the embodied (scarce) net input is likely a more important indicator. For example, in the marine ecosystem (Table 1) the export (fishing) of 0.06 units of "large fish" carried with it more embodied solar energy than 4 units of pelagic fish. This sort of conclusion should be of great importance to ecosystem managers.

The direct and indirect impact on each process in the marine ecosystem of exporting an additional unit of large fish is demonstrated. The energy intensity is the sum of the extra respiration energy, on the part of *all* the processes, needed to carry out the extra direct and indirect production flows, which in turn were necessary for the increased unit of exported large fish. The added respiration and production flows needed from each process to support the extra unit of export are also calculated.

The system of equations is also applied to an oyster reef ecosystem where large amounts of indirect self-use are demonstrated. This system has a high density of feedback loops, which accounts for the self-use phenomenon.

The final application of the equations was on data from disturbed and undisturbed tidal marshes in Florida. The resulting intensities clearly showed which of the 17 processes in this complex system were most damaged by the heating of the marsh ecosystem. They also showed that *all* processes were distressed to some extent, and they revealed the hierarchy of solar energy dependence under disturbed and undisturbed conditions.

In general, this paper describes both an accounting system and an elementary theory of value determination. It was applied to the three very different ecosystems with demonstrable results. It seems that this approach should be useful to ecosystem planners and managers, biological oceanographers, as well as ecologists studying laboratory systems.

References

BARBER, M. 1978. A retrospective Markovian model for ecosystem resource flow. Ecol. Model. 5: 125–135.

COSTANZA, R., AND C. NEILL. 1981. The energy embodied in the products of ecological systems: a linear programming approach. p. 661–670. *In* R. Mitsch [ed.] Energy and Ecological Modeling Conference. University of Louisville, Louisville, KY.

DAME, R., AND B. PATTEN. 1981. Analysis of energy flows in an intertidal oyster reef. Mar. Ecol. Prog. Ser. 5: 115–124.

FINN, J. 1976. Measure of ecosystem structure and function derived from the analysis of flows. J. Theor. Biol. 56: 363–380.

HANNON, B. 1973. The structure of ecosystems. J. Theor. Biol. 41: 535–546.

1976. Marginal product pricing in ecosystems. J. Theor. Biol. 56: 256–267.

1979. Total energy costs in ecosystems. J. Theor. Biol. 80: 271–293.

1982. Discounting in ecosystems. Presented at Wallenberg Symposium on Economics and Ecology, Stockholm, Sept. 1982.

HANNON, B., R. COSTANZA, AND R. HERENDEEN. 1983. Measure of energy cost and value in ecosystems. J. Environ. Econ. Manage. (In press)

HERENDEEN, R. 1981. Energy intensities in ecological and economic systems. J. Theor. Biol. 91: 607–620.

HOMER, M., AND W. KEMP. 1983. Unpublished manuscript. See Ulanowicz (1983) for data and description.

LEVINE, S. 1977. Exploitation interactions and the structure of ecosystems. J. Theor. Biol. 69: 345–355.

PATTEN, B., R. BOSSERMAN, J. FINN, AND W. CALE. 1976. Propagation and cause in ecosystems, p. 457–579. *In* B. C. Patten [ed.] Systems analysis and simulation in ecology. Vol. 4. Academic Press, New York, NY.

PETERS, R. 1983. Ecological implications of body size relationships. Cambridge University Press, London.

SAMUELSON, P. A. 1966. A new theorem on non-substitution. The Collected Papers of P.A.S., J. E. Stiglitz [ed.] Vol. I, Chapter 37, MIT press, Cambridge, MA.

STEELE, J. 1974. The structure of marine ecosystems. Harvard Univ. Press, Cambridge, MA., page 20: as interpreted in Ulanowicz and Kemp (1979).

ULANOWICZ, R. 1983. Identifying the structure of cycling in ecosystems. Math. Biosci. 65: 219–237.

1984. Community measures of marine food networks and their possible applications, p. 23–47. *In* J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems. Plenum Publishing Corp., NY.

Appendix A

Deriving the Commodity Weightings or Intensities for Processes with More Than One Kind of Commodity Output

Let U be a matrix such that u_{ij} is the amount of commodity i (row) *used* by process j. Let V be a matrix such that v_{jk} is the amount of commodity k (column) *made* by process j. Let vector q be the sums down the columns of V (the total commodity output) and g be the sums across the rows of V (the total process output). The vectors e and r are the same as defined in the main text, the vector of the scarce net input and the net output vector, respectively.

To find the total commodity output, one begins with, $q = Uu + r$, where u is the vector of ones, or

$$q = UV^{T-1}V^{T}u + r,$$

where T means matrix transpose and -1, matrix inversion, or $= UV^{T-1}q + r$, definition of q, or

(A1) $\quad q = (I - UV^{T-1})^{-1}r.$

The balance of the net input across the processes is,

(A2) $\quad \epsilon U + e = \epsilon V^{T},$

or $\epsilon = e(V^{T} - U)^{-1}$, or

(A3) $\quad \epsilon = eV^{T-1}(I - UV^{T-1})^{-1}.$

Here V^{T-1} plays the role of $\hat{P}^{-1}$, and U is the equivalent of P. Consequently, the commodity vector obtained from the sum down the columns of V^{T} should agree with the sum across the rows of $U + R$. In other words, V does not contain the non-basal metabolism flows. The vector ϵ is the intensity of the *commodities* in terms of the scarce medium.

The physical constraints require that $\epsilon > 0$, but equation (A3) does not guarantee this inequality, because there is no reason why all the elements of V^{T-1} must be positive. However, not all of these elements *need* to be positive in order that $\epsilon > 0$. Experience with economic data indicates that some of the ϵ become negative (or zero) whenever the aggregation procedure is improper. With highly detailed systems, the accuracy of the data varies from commodity to commodity. The commodities about which little accurate data are available are sometimes lumped together. This procedure occasionally results in some negative components of ϵ. By careful reconsideration of the aggregation techniques, the problem can be avoided.

Note that in equations (A1) and (A3), U and V must be square. This squaring process is also somewhat arbitrary (as it is with the formation of P). There are generally more commodities than processes. The combination of similar biomass or fecal products into aggregate commodities, for example, seems to be a reasonable procedure for equating the number of commodities to the number of processes.

In the "Use-Make" formulation, the basal metabolism should be considered as a part of the net output of the commodity (e.g., algal biomass) with which it is associated, since this metabolism is supposed to represent a depreciation of that commodity. The same procedure was required in the single commodity formulation.

Certain economic formulations of the U and V do not require squareness of these matrices in order to formulate the ϵ. However, they require normalization by the row sums

of V (total process outputs). Yet these sums (g_i) are meaningless in most physical systems. They exist in economic systems only when the commodities have been weighted by a set of system prices. If we weight the commodities with the ϵ (at this point unknown) the solution for the ϵ becomes identical to equation (A3) (Hannon et al. 1983), requiring squared U and V matrices. In a theoretical sense, the approach outlined here is certainly more desirable than the simpler method outlined in the text. Consider, for example, the flows in the North Sea ecosystem (Fig. 2) between the Herbivores and processes 3, 4, and 5. Although the approach given earlier requires these three flows to be identical in composition, they obviously are not. Flows to processes 3 and 4 are probably quite similar, but flow 2−5 is a fecal output from the herbivore to the bacteria. Likewise, flows 6−8 and 6−9 are probably not the same. Obviously, more data from the system is required before the "Use-Make" approach can be applied, but the additional information in the results should make the extra effort worthwhile.

Appendix B

Proof that the Intensities Vary from Unity Depending on the Level of the Non-Basal Metabolism

Note that the column sums of P give,

(B1) $\quad uP + e = P + w,$

where w is the non-basal metabolism, u is a vector of ones, and P, r, and e are assumed to all be measured in the same units.

Multiplying equation (B1) from the right by $\hat{P}^{-1}$, the diagonalized inverse of P gives,

(B2) $\quad uG + e\hat{P}^{-1} = u + w\hat{P}^{-1},$

or

$$uG + \epsilon(I - G) = u + w\hat{P}^{-1},$$

or

$$\epsilon = u + w\hat{P}^{-1}(I - G)^{-1}.$$

Therefore, if $w = 0$, then $\epsilon = u$, the unit vector; that is, if all of the respiration is included in r, the intensities are all ones. The non-basal metabolism generates the variation in the intensities. This phenomenon is masked when the units of measurement are not the same throughout. In the case of mixed units in P, u becomes the vector which transforms the various units of the commodities into the units of e. Equation (B2) still holds in such a situation, and ϵ would equal the transformation vector when $w = 0$.

Appendix C

Finding the Total Input Increases for Each Component When the Net Output is Increased

Equation (B2) is useful for formally calculating the change in total input to each process under a given change in net output, Δr. Rearranging equation (B2),

(C1) $\quad w = (\epsilon - u)(I - G)\hat{P}.$

Then

$$w + \Delta w = (\epsilon - u)(I - G)(\hat{P} + \Delta\hat{P}), \text{ or}$$

(C2) $\Delta w = (\epsilon - u)(I - G)\Delta\hat{P}$,

(C3) $= w\hat{P}^{-1}\Delta\hat{P}$,

as was described verbally in the text, and this result is independent of the variation in commodity units. The total input change, ΔQ, is the change in output plus the change in non-basal metabolism:

(C4)
$$\Delta Q = \Delta w + \Delta P, \text{ or}$$
$$\Delta Q = \Delta r(I - G)^{T-1}(wP^{-1} + u) = \Delta rH,$$

where superscript T means "transpose" of $(I - G)^{-1}$.

For example, in the North Sea ecosystem Δr is taken as a row vector of zeros except that the tenth entry becomes 1.0. When the specification of the change in desired net output, Δr, is this simple, it is easier to calculate ΔQ from ΔP and equation (C3). However, if the specification is complex (e.g., multiple exports), then the matrix H in equation (C4) should be calculated. For Steele's ecosystem that matrix is (in rounded-off form):

	1	2	3	4	5	6	7	8	9	10
1	10	0	0	0	0	0	0	0	0	0
2	19.2	1.92	0	0	0	0	0	0	0	0
3	148	14.8	7.73	0	0	0	0	0	0	0
4	408	40.8	10.6	12	0	0	0	0	0	0
5	47.5	4.75	0	0	2.48	0	0	0	0	0
6	115	11.5	0	0	6	2.4	.42	0	0	0
7	49.9	4.98	0	0	2.6	0	1.05	0	0	0
8	2840	284	16.4	18.5	115	46.1	8.07	13.9	7.7	0
9	1150	115	0	0	60	24.0	4.2	0	10	0
10	28400	2840	164	185	1150	462	80.7	139	77	10

The entries in the tenth row are the increases in total inputs required for an increase of one unit in the export of large fish. The fourth row shows the much lower total inputs needed for a unit increase in the exports of pelagic fish. Note that the first column of the H matrix is the vector of energy intensities, ϵ. These intensities, therefore, represent the total net inputs to the primary producer in order that the respective processes can have one unit of net output.

The calculation of increase in total input is straightforward whenever no basal metabolism is included in the net output (as demonstrated in the text and above). However, the presence of the basal metabolism vector in the net output and the correlation of basal rates with the stock size of each process requires that the total input increase vector be calculated in an iterative fashion. Here is the stepwise procedure:

1) Specify Δr, the desired export amounts.
2) $\Delta P = (I - G)^{-1}\Delta r$, equation (3) in incremental form.
3) $\Delta w = w\hat{P}^{-1}$, ΔP assuming $w\hat{P}^{-1}$ are constants.
4) $f(r + w, \Delta(r + w)) = r_m + \Delta r_m$, the needed vector of concave functional relationships for each component between respiration and basal metabolism.
5) Respecify $\Delta r = \Delta r + \Delta r_m$ and substitute this into (2) until the resultant change in P becomes desirably small.
6) The final $\Delta P + \Delta w$ is the desired result.

Note that this iterative process may not involve specific statements of the actual increases

in the stocks. We must realize, however, that the above iterative process is not duplicating the actual growth process. It is a process which searches for the new steady-state (where the stock increases or decreases are zero). We arrive at this steady-state via a set of assumptions: G = constant; $w\hat{P}^{-1}$ = constant and f is concave. These all appear to be reasonable for small Δr.

Appendix D

Establishing Time Value in Ecosystems

The analogy of energy intensity to price or measure of value is based on the assumption that the age of a specified unit of biomass, for example, has no effect on its exchange value. This is almost never the case in economics, where the price of a unit of production is affected by the age of its inputs, by its production techniques, and by the level of demand. The age-based weighting is done with an "interest" rate, a number which gives the unit charge to the borrower for a given time period. A lender is given "interest" in payment for forgoing consumption and for assuming the risk that repayment might never be fully made.

In the ecological context, the economist might note that the scarce input, say the resource solar energy, could be used in immediate consumption, "invested" in biocapital, or some combination of both. The economist then looks for that combination of direct and indirect resource uses (through investment) which minimizes the total use of the resource for a given net output of the system. For example, if fossil energy were deemed scarce, then the ratio of such energy input per dollar of net output might be minimized in a human economy. This is the nature of the efficiency criterion in economics.

In economics, the resource is valued in terms of a system currency, therefore, the choice of direct−indirect balance involves the concept of pure time value. Investment in biocapital requires a waiting period between the moment of investment and the moment of output. A surplus of output over input must be provided, and its magnitude depends on the length of the waiting period and the level of need for immediate consumption of the resource. This surplus is commonly called an interest payment. Consequently, the initial cost of biocapital investment must be augmented by an interest payment, which is cast in terms of the resource input. Biocapital investment so weighted may enable the system to provide the net output using the fewest resources, up to a given investment level — beyond which direct resource consumption becomes more efficient (or conversely). So, in our case, we assume that the ecosystem strives for resource efficiency (the most net output value per unit resource used). The ecosystem strikes a balance between direct resource consumption (reproduction, predation, and predator avoidance) and investment (growth of the biomass). The important thing to realize is that waiting time has a measurable cost, or interest, which may be cast in resource terms. This interest cost is a function of the waiting time, the production techniques available and the "impatience" for direct resource consumption. Thus, there is a tradeoff between immediate and delayed uses of the resource in terms of net output production. We might assume that the ecosystem maximizes its "profit", or net output, per unit of resource available, or perhaps better, that it maximizes the *value* of the profit (ϵr).

In the most general sense, the economist assumes that the *value* of the marginal[7] profit (marginal[7] price less marginal[7] cost) is the same in each period. This equality is accomplished by assuming an interest rate. The assumption leads to the result that marginal proportional profit rate is equal to the interest rate. Both the ability to produce and the desire to consume the resource are incorporated in this result.

[7]Cost of the last unit produced, as opposed to the average cost of all the units ever produced.

For a more restricted case, Samuelson (1966) has shown that if the production technique is describable by a function which is homogenous of degree one, and has only one type of net input, then unit value or price is determinable solely from a knowledge of production and the interest rate alone — independent of the demand mix or level. In this paper I meet these conditions with the constant $\boldsymbol{G}$ and $\boldsymbol{E}\hat{\boldsymbol{P}}^{-1}$ (or their more elaborate version in Appendix A), except that the interest rate was assumed zero.

Biological interest rates are not posted in the system as they are in economic systems. What experiments might be done to find them? I can imagine several. Consider the pine needle on a conifer. It is well known that the needle stays on the branch for several seasons, its ability to fix light declining each period. So the tree makes an initial "investment" in a needle, and in turn it captures energy (net, after maintenance) in declining amounts for a number of seasons. The cumulative capture can be compared with the initial investment (as an efficiency test), only if the capture in each period is translated to the initial time by means of an *interest factor* (see below). This factor, if less than one, is an acknowledgement of the cost to the tree of waiting for a return after its initial investment. The tree could not grow if the cumulative net production by the needle were less than or equal to the initial investment. A surplus is required.[8] This surplus is evidence of the existence of an interest factor. If the present-valued net production during period i is P_i is equal to the initial investment, then the present value of the net production in the nth period becomes [9]$P_n(1 + q)^n$. Furthermore, if the initial cost of producing the needle is I_o, then $I_o = \Sigma_n P_n(1 + q)^n$. Since (in theory) the I_o and P_n are known, this equation can be solved for the biological interest rate on conifer needles. A similar argument can be made for the leaves on a deciduous tree. The tree will add leaves in the understory until the energetic cost of the specific (marginal) additional leaf is equal to the *present value* of the future production energy (net of any maintenance cost) of the added leaf. The same argument could be made for algae in a pond, lake or ocean. The deeper algae are added only if their energetic cost is less than or equal to less than the present value of their future net energy production.

The experimental problems here seem formidable to this theoretician. How is the energy cost of a leaf or a needle to be approximated? Could an input−output analysis as described in the text or Appendix A be done on a tree to obtain the q for a given set of understory leaves? How is the net energy production of the dropped needle or the new marginal leaf to be calculated? Could these calculations be approximated from existing data?

The concept of biological interest also arises at the level of the biosphere. Does life on earth retard or increase the rate of entropy production in the universe? My view is that the entropy formation rate might be retarded by the presence of life because of the retention time of the captured energy. In the steady state biosphere, energy is absorbed and later released. But the energy flows are always balanced. The entropy production at each instant (the difference between the high quality energy absorbed and the heat radiated) is balanced (I speculate) by the effect of the retention of the high quality energy in the biosphere and the presence of biological interest rates in the ecosystem. The captured high quality energy is retained in the steady state system just long enough to offset the rate of entropy formation. The entropy input rate is increased (by the factor e^{qT}, where T is the mass-weighted average life-time of all organisms) to the present and is equal to the entropy output rate. This speculation reveals the interest rate of the biosphere (q), and here economics and thermodynamics touch in a satisfying and seemingly consistent way.

---

[8]The surplus in the steady state condition is the depreciation. Thus, a non-zero discount factor exists for the steady state under the view taken in this Appendix.

[9]$(1 + q)^n$ is the interest factor.

Energy Cycling, Length of Food Chains, and Direct versus Indirect Effects in Ecosystems[1]

BERNARD C. PATTEN

Department of Zoology and Institute of Ecology, University of Georgia, Athens, Georgia 30602, USA

Introduction

It is difficult to unlearn things learned. Traditional ecology teaches a number of principles to its students, for example, that energy does not cycle in ecosystems, although matter does; that food chains are short because the number of trophic transfers possible is thermodynamically limited; that direct biotic interactions, such as competition and predation, are important mediators of adaptations for evolutionary fitness. Another model of these same phenomena, a systems model, suggests a different set of conclusions. The purpose of this paper will be to demonstrate these new viewpoints and, implicitly, to argue for the strong introduction of systems analysis into biological oceanography. The points will be made by analyzing a small marine ecosystem model of energy flow.

Model Description

Figure 1 depicts a compartment model for energy flow (kcal m^{-2} d^{-1}) and storage (kcal m^{-2}) in an intertidal oyster reef community. The filter feeding compartment (1) consists of the American oyster (*Crassostrea virginica*) and, secondarily, a mussel (*Brachidontes exustus*). The shells of these bivalve mollusks form the frame and inner spaces that create habitat and, in addition, the animals filter particles suspended in the ambient water around them. These particles form the trophic base for the community. Deposited detritus (2) comprises the feces and pseudofeces produced by the feeding and excreting mechanisms of the shellfish. Microbiota (3) consists of bacteria, yeasts and fungi associated with detritus. The meiofauna (4) is defined as benthic animals that pass through a 1 mm sieve, but are retained by a 0.063 mm sieve. Deposit feeders (5) consist of macrofauna that feed in the sediments. Predators (6) are the animals directly benefitting from this mode of interaction (predator–prey) with other forms. The energy flow processes that internally couple these compartments, and also those that provide input and output linkage to environment, are detailed in Dame and Patten (1981), together with documentation of the numerical data. The single system input (z_1) consists of phytoplankton and suspended detrital particles acquired by filter feeding. Output processes include respiration (y_k, $k = 1, 3, 4, 5, 6$), mortality (y_1, y_6) and resuspension (y_2). The intrasystem flows, f_{ij} (from compartments j to i), are realized through feeding and egestion interactions between the different compartments, $i, j = 1, \ldots, n$, where $n = 6$.

Dynamic equations for this system can be written as,

$$\dot{x}_j = \sum_{\substack{k=1 \\ k \neq j}}^{n} f_{jk} + f_{jj} + z_j - \sum_{\substack{i=1 \\ i \neq j}}^{n} f_{ij} - f_{jj} - y_j, \qquad j = 1, \ldots, n,$$

[1]University of Georgia, *Contributions in Systems Ecology*, No. 66 and *Okefenokee Ecosystem Investigations, Theoretical Series*, No. 2.

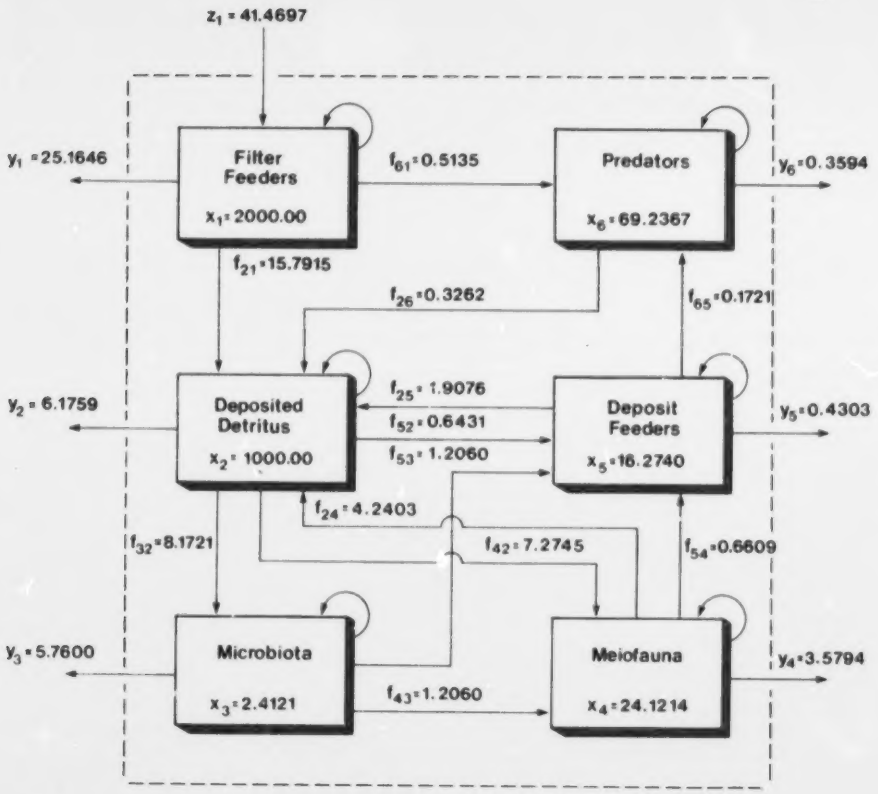


FIG. 1. Intertidal oyster reef compartment model. Numbers within the rectangles indicate steady state standing crop energy storages (x_i, $i = 1, \ldots, 6$, in kcal m^{-2}), and those associated with arrows denote energy flows (input z_i, outputs y_i, and internal f_{ij}, $i, j = 1, \ldots, 6$, in kcal m^{-2} d^{-1}).

where f_{ij} and f_{jk} are intercompartmental flows, z_i and y_i are inflows and outflows, respectively, and f_{ii} are flow contributions to and from storage, x_i. The terms f_{ii} can be cancelled or left in without altering any dynamic properties. At steady state $\dot{x}_i = 0$, and two types of total flow, or throughflow, may be defined:

$$\sum_{\substack{k=1 \\ k \neq j}}^{n} f_{jk} + f_{jj} + z_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} f_{ij} + f_{jj} + y_j \triangleq T_j,$$

and

$$\sum_{\substack{k=1 \\ k \neq j}}^{n} f_{jk} + z_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} f_{ij} + y_j \triangleq \hat{T}_j.$$

The first is termed *inclusive throughflow* (Higashi and Patten 1984), meaning that the contribution to storage f_{jj} is considered; this is the basis of environ analysis (Matis and

Patten 1981). The second, based on cancellation of the f_{ij}'s in the first equation above, is the conventional *throughflow* of economic input-output analysis (e.g., Leontief 1966; Hannon, this volume) and ecological flow analysis (e.g., Hannon 1973; Patten et al. 1976; Finn 1976). These methods will be used subsequently to investigate energy cycling, food chain relationships, and direct vs. indirect effects in the oyster reef model.

Energy Cycling

Adjacency Matrix

The digraph (directed graph), which Fig. 1 in effect represents, is isomorphic to an adjacency matrix $A = (a_{ij})$, where $a_{ij} = 1$ denotes a direct energy flow from compartment j to i, and $a_{ij} = 0$ signifies the absence of such an energy connection; $a_{jj} = 1$ defines an energy storage in j, and $a_{jj} = 0$ denotes no storage. Table 1a shows the adjacency matrix for Fig. 1; $a_{jj} = 1$ for all $j = 1, \ldots, 6$ because all compartments in the oyster model store energy.

Markov Transition Matrix

The energy flows $F = (f_{ij})$ in Fig. 1 can be normalized or nondimensionalized by computing another matrix (f_{ij}/x_j), x_j the storage in compartment j, and then making the transformation $P = I + h(f_{ij}/x_j)$, I the identity matrix of proper order, and h a time scale

TABLE 1. Oyster reef model first order matrices: (a) A for paths; (b) P for normalized energy flows (for example, $9.948 - 1 = 9.948 \times 10^{-1}$).

	Compartments						
			(a)				
From	1	2	3	4	5	6	Row Sum
To							
1	1	0	0	0	0	0	1
2	1	1	0	1	1	1	5
3	0	1	1	0	0	0	2
4	0	1	1	1	0	0	3
5	0	1	1	1	1	0	4
6	1	0	0	0	1	1	3
Column Sum	3	4	3	3	3	2	18
			(b)				
From	1	2	3	4	5	6	Row Sum
To							
1	9.948−1	0	0	0	0	0	9.948−1
2	1.974−3	9.944−1	0	4.395−2	2.930−2	1.178−3	1.071
3	0	2.043−3	1.530−1	0	0	0	1.551−1
4	0	1.818−3	1.250−1	9.121−1	0	0	1.039
5	0	1.608−4	1.250−1	6.850−3	9.614−1	0	1.093
6	6.419−5	0	0	0	2.644−3	9.975−1	1.000
Column Sum	9.969−1	9.985−1	4.030−1	9.629−1	9.934−1	9.987−1	5.353

factor. Here, $h = 0.25$d was used. In F, the diagonal elements are summed outflows from each compartment, $f_{jj} = -\sum_{i=0}^{n} f_{ij}$, where $i = 0$ is the system environment. The scalar h is selected such that $0 \leqq h(f_{ij}/x_j) \leqq 1$ when $i \neq j$, and $0 \leqq 1 - h\sum_{i=0}^{n} f_{ij} \leqq 1$ when $i = j$. Thus, $0 \leqq p_{ij} \leqq 1$, where $(p_{ij}) = P$, and the p_{ij}'s are probabilities, defined in effect as $p_{ij} = f_{ij}/T_j$. P is then a one-step transition matrix for a discrete time Markov chain (Kemeny and Snell 1960) with time step h. P is shown in Table 1b for the Fig. 1 model. Its entries define the 6-hourly fractions of initial nondimensional units ($P^0 = I$) of energy in each compartment j transferred from j to each i. The diagonal elements, $p_{jj} = 1 - h\sum_{i=0}^{n} f_{ij}/x_j$, denote energy fractions not transferred out of j in the transition interval, i.e., storage. A second Markov chain, $\hat{P} = (\hat{p}_{ij}) = f_{ij}/\hat{T}_j$, can also be defined based on conventional throughflow, which excludes the consideration of storage. The diagonal elements of this matrix are therefore zero ($\hat{p}_{jj} = 0$). It will also be useful in later sections to consider the matrix P with its diagonal entries zeroed. This matrix will be denoted $P(0)$, and is not to be confused with $\hat{P}$, from which it is distinct except for the common possession of zero diagonal elements.

Cycles and Storages

Returning to the adjacency matrix A, A^k gives the number of paths of length k from each j to each i in the system (Roberts 1976, p. 54). When $i = j$, $a_{jj}^{(k)}$ denotes the number of cycles of length k associated with compartment j. Table 2a shows the numbers of cycles of selected lengths k in the Fig. 1 model. The most striking feature of these numbers is how large they become at the longer path lengths. This is because the series $\sum_{k=0}^{\infty} A^k$ is divergent and thus $\lim_{k\to\infty} a_{ij}^{(k)} = \infty$.

A cycle is simple if it contains no repeated compartments. The maximum length of a simple cycle in an n'th order system is n. Therefore, the maximum length of a simple cycle is 6 for the Fig. 1 model; the longest actual cycles are only length $k = 5$, however, because no digraph arc (arrow) exists from compartment 6 to 1. Ulanowicz (pers. comm.) has enumerated and identified the simple cycles in this model using a network analysis program (Ulanowicz 1982). There are 33 of them, as listed in Table 3, but since many are redundant (for example, $2 \rightarrow 5 \rightarrow 6 \rightarrow 2$, $5 \rightarrow 6 \rightarrow 2 \rightarrow 5$ and $6 \rightarrow 2 \rightarrow 5 \rightarrow 6$ are all the same cycle) only 10 of the 33 are unique; these are listed under compartment 2.

Simple cycles of length k, with every intermediary compartment between the originating and terminal compartments different, represent one end of the spectrum of cycles. The other, with no intermediary compartment different from the beginning and end compartments, represents k'th order storage, that is, storage for kh units of time. All other length k cycles are compound true cycles, and they may have intermediary cycles, nested cycles, and storages of path lengths less than k in various combinations. The vast majority of cycles in the Fig. 1 model are of this type, since there are only 33 simple cycles and, for each cycle length k, there is only one storage.

The question now is, do these cycles inherent in the structure of the oyster model really carry any energy?

Higher Order Cycling

Higher powers P^k of P represent the k'th order distribution of an initial unit of energy, $P^0 = I$, in each compartment to all compartments after k transitions. That is, $p_{ij}^{(k)}$ denotes the fraction of the original energy, $p_{jj}^{(0)} = 1$, in compartment j that is transferred to compartment i over paths of length k in an amount of time hk. Because energy is lost from the system ($p_{0j} > 0$ for all j in Fig. 1), the system is dissipative. This is the basis (Patten

TABLE 2. (a) Numbers of cycles of lengths $k = 1, 2, 3, 10, 50$ associated with each compartment of the Fig. 1 model. (b) Nondimensional energy flows around these cycles.

(a)

Compartments	Path lengths, k				
	1	2	3	10	50
1	1	1	1	1	1
2	1	3	11	34729	2.894+24
3	1	1	3	11032	9.194+23
4	1	2	6	18560	1.547+24
5	1	2	7	23696	1.975+24
6	1	1	2	7527	6.273+23

(b)

	1	2	3	10	50
1	9.948−1	9.897−1	9.845−1	9.494−1	7.712−1
2	9.944−1	9.890−1	9.837−1	9.494−1	7.996−1
3	1.530−1	2.342−2	3.602−3	1.581−4	3.622−4
4	9.614−1	9.244−1	8.887−1	6.753−1	1.460−1
5	9.614−1	9.244−1	8.887−1	6.753−1	1.460−1
6	9.975−1	9.951−1	9.926−1	9.755−1	8.835−1

	100	500	2000
1	5.947−1	7.438−2	3.060−5
2	6.555−1	1.373−1	5.157−4
3	3.336−4	7.412−5	4.167−7
4	9.819−3	2.069−3	8.409−6
5	2.833−2	2.270−3	1.474−5
6	7.806−1	2.910−1	7.126−4

et al. 1976; Patten 1978) for the series $\sum_{k=0}^{\infty} P^k$ to converge to another matrix, $Q = (I - P)^{-1}$. The elements q_{ij} of this latter "transitive closure" matrix represent the time (or path) integrated distribution of nondimensional original energy (one unit in each compartment) to all other compartments over all paths of all lengths (i.e., over all future time) until dissipation is complete. The terms P^k represent a partition of the integral energy Q according to paths A^k of length k, and thus the energy associated with paths of different lengths can be quantified. Analogous developments are possible from the $\hat{P}$ matrix, i.e., $\sum_{k=0}^{\infty} \hat{P}^k = (I - \hat{P})^{-1} = \hat{Q}$.

Table 2b denotes the energy flows over cycles of selected lengths from $k = 1$ to $k = 2000$. Reading from left to right, each row quantifies the dissipation rate of that compartment. Compartment 3 (microbes) has the most rapid energy turnover, and compartment 6 (predators) the least. Compartment 2 (detritus) initially stores slightly less ($p_{22} = 9.944 \times 10^{-1}$) than compartment 1 (bivalves) does ($p_{11} = 9.948 \times 10^{-1}$) during the first transition step (i.e., loses more); nevertheless, by virtue of its intrasystem coupling it has an order of magnitude more energy remaining at the end of the interval ($p_{11}^{(2000)} = 3.060 \times 10^{-5}$, $p_{22}^{(2000)} = 5.157 \times 10^{-4}$). It is clear since all the initial energy in the system is not dissipated by $k = 2000$ ($hk = 500$ days), that the long cycles counted in great numbers in Table 2a are indeed utilized in energy cycling in this model. But how much of the total energy flow actually is around cycles?

TABLE 3. Simple cycles of the Fig. 1 model, listed by originating compartments. The unique cycles are the ten listed under compartment 2.

Compartment 1	*Compartment 2*	*Compartment 3*
none	2→3→4→2	3→4→2→3
	2→3→4→5→2	3→4→5→2→3
	2→3→4→5→6→2	3→4→5→6→2→3
	2→3→5→2	3→5→2→3
	2→3→5→6→2	3→5→6→2→3
	2→4→2	
	2→4→5→2	
	2→4→5→6→2	
	2→5→2	
	2→5→6→2	
Compartment 4	*Compartment 5*	*Compartment 6*
4→2→3→4	5→2→3→4→5	6→2→3→4→5→6
4→2→4	5→2→3→5	6→2→3→5→6
4→5→2→3→4	5→2→4→5	6→2→4→5→6
4→5→2→4	5→2→5	6→2→5→6
4→5→6→2→3→4	5→6→2→3→4→5	
4→5→6→2→4	5→6→2→3→5	
	5→6→2→4→5	
	5→6→2→5	

CYCLING INDEX

Based on conventional throughflow, Finn [1978, equation (29)] formulated the ratio of k'th order cycled to total flow in a system as a cycling index,

$$\hat{\mathrm{CI}}(k) = \sum_{j=1}^{n} \hat{\mathrm{RE}}_j(k)\hat{T}_j(k)/\hat{T}(k),$$

where $\hat{\mathrm{RE}}_j(k)$, recycling efficiency, is the ratio of recycled to noncycled substance flowing through compartment j over cycles of length k, and $\hat{T}(k) = \sum_{j=1}^{n} \hat{T}_j(k)$ is the k'th order total system throughflow. The global recycling efficiency $\hat{\mathrm{RE}}_j$, defined for all $k \geqq 0$, was formulated in terms of the diagonal elements $\hat{q}_{jj}$ of $\hat{Q}$ [Finn 1978, equation (26)]:

$$\hat{\mathrm{RE}}_j = \frac{\hat{q}_{jj} - 1}{\hat{q}_{jj}} = \frac{\hat{q}_{jj} - \hat{p}_{jj}^{\,0}}{\hat{q}_{jj}} = 1 - \frac{\hat{p}_{jj}^{\,0}}{\hat{q}_{jj}}.$$

Here, $\hat{p}_{jj}^{\,0} = 1$, the initial inflow in compartment j, which is properly not counted as cycled substance. The integral cycling index for all $k \geqq 0$ is then

$$\hat{\mathrm{CI}} = \sum_{j=1}^{n} \hat{\mathrm{RE}}_j \hat{T}_j / \hat{T},$$

whose value is 0.104 for the Fig. 1 model.

CYCLING INDEX CORRECTION FOR STORAGE

If this scheme were automatically applied to the inclusive throughflow case, then the corresponding expression for the cycling index would be,

$$CI = \sum_{j=1}^{n} Re_j T_j / T,$$

where

$$RE_j = \frac{q_{jj} - 1}{q_{jj}}.$$

This would give a value of CI = 0.913 for the oyster model, and thus most of the energy flow would be indicated to be cycled. The recycling efficiencies RE_j, $j = 1, \ldots, 6$, contributing to this high cycling index are 0.995, 0.996, 0.231, 0.927, 0.964 and 0.998. The value $RE_1 = 0.995$ illustrates a problem however: compartment 1 can cycle no energy because it contains no incoming arcs from other compartments (Fig. 1). The number 0.995 is really the one-step transition probability (Table 1b), and is due entirely to storage, not cycling.

It was pointed out above that one of the $a_{jj}^{(k)}$ k'th order cycles represented a path for pure storage, of the form $j \rightarrow j \rightarrow \ldots$ ($k - 2$ terms) $\rightarrow j$. The "flow" over this path corresponds to the amount of storage, and is computed by p_{jj}^{k}. The higher order storages must be subtracted from the RE_j values above to obtain true recycling,

$$RE_j = \frac{q_{jj} - \sum_{k=0}^{\infty} p_{jj}^{k}}{q_{jj}} = 1 - \frac{\sum_{k=0}^{\infty} p_{jj}^{k}}{q_{jj}} = 1 - \frac{1}{q_{jj}(1 - p_{jj})},$$

and then the cycling index, CI, above applies (Patten 1984a; Patten and Higashi 1984). The recycling values are now lower, $RE_j = 0$, 0.280, 0.092, 0.229, 0.110 and 0.005, respectively, for $j = 1, \ldots, 6$, and hence also the cycling index, CI = 0.110. Eleven percent of the energy flow through the oyster reef model is cyclical vs. 10% in the conventional case, $\hat{CI} = 0.104$.

Time Development of Cycling

At path length k, the ratio of cycled to cycled plus stored material at each compartment j is,

$$RE_j(k) = \frac{p_{jj}^{(k)} - p_{jj}^{k}}{p_{jj}^{(k)}} = 1 - \frac{p_{jj}^{k}}{p_{jj}^{(k)}},$$

where $p_{jj}^{(k)}$ is the j'th diagonal element of P^k, and p_{jj}^{k} is the j'th diagonal element of P raised to the k'th power. Table 4 provides values for $RE_j(k)$ for selected cycle lengths k in the oyster model. At small path lengths there is little cycling and mostly storage. Cycling develops at different rates for different compartments with increasing path length. The boldface entries identify where cycling first exceeds 10% of the total cycling plus storage. This occurs at $k = 5$ for compartment 3, $k = 20$ for compartment 4, $k = 100$ for compartments 2 and 5, and never happens with compartments 1 and 6.

Storage dominates cycling at small path lengths, because for storage paths the component probabilities p_{ij}, being diagonal elements, are higher than the nondiagonal probabilities, p_{ij}, $i \neq j$ (Table 1b), that typify cyclic paths. Furthermore, with increasing path length storage paths will dissipate more slowly than cyclic paths because of these higher probabilites. However, at any given path length the number of storage paths remains one, whereas the number of cycles, $a_{jj}^{(k)} - 1$, increases combinatorially. Therefore, while the contributions of each cyclic path may be small compared to that of the storage path, at

TABLE 4. Recycling efficiencies at paths of selected lengths k, $RE_i(k)$, for the six oyster model compartments. The four boldfaced entries indicate first places in the sequences where the ratios reach at least 10%.

Cycle lengths, k	Compartments					
	1	2	3	4	5	6
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	8.558−5	0	9.606−5	5.097−6	0
3	0	2.692−4	5.194−3	3.121−4	2.430−5	5.044−10
4	0	5.467−4	7.085−2	6.617−4	5.987−5	2.840−9
5	0	9.115−4	**4.350−1**	1.158−3	1.130−4	8.446−9
7	0	1.879−3	9.818−1	2.645−3	2.763−4	3.521−8
10	0	3.840−3	1.000	6.349−3	6.794−4	1.352−7
20	0	1.347−2	1.000	**3.832−1**	3.758−3	1.406−6
50	0	5.395−2	1.000	5.154−1	4.126−2	2.204−5
100	0	**1.270−1**	1.000	9.897−1	**3.082−1**	1.397−4
200	0	2.604−1	1.000	1.000	9.476−1	7.172−4
500	0	5.532−1	1.000	1.000	1.000	4.726−3
1000	0	8.117−1	1.000	1.000	1.000	1.630−2
1500	0	9.247−1	1.000	1.000	1.000	3.097−2
2000	0	9.725−1	1.000	1.000	1.000	4.692−2

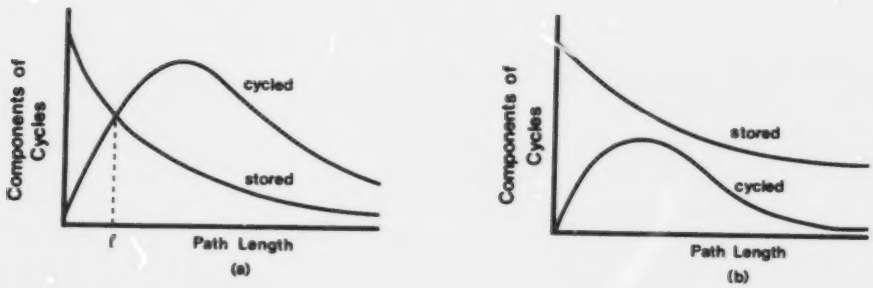


FIG. 2. Two cases of cycled vs. stored components of energy flow over cyclic paths of different lengths. (a) Cycled component exceeds stored component beginning at path length ℓ. (b) Cycled component never exceeds stored component with increasing path length.

some path length the profusion of cycles may produce a summed quantity of cycled material that exceeds the storage. In the oyster model this happened for all compartments except 1 and 6 by path length 2000 (Table 4). The relationships just described are schematized in Fig. 2; diagram (a) applies to compartments 2−5 and (b) to 6.

Hypothesis 1. The general principle that energy does not cycle in ecosystems is false. Energy, like matter, does cycle in the ecosystem.

Food Chain Length and Number of Trophic Levels

Energy Transfer Efficiency

Lindeman (1942) investigated "food cycles," and was first to point out the inefficiency of energy transfer between trophic levels. This became widely accepted and promoted as a thermodynamic explanation for Elton's (1927) pyramids of animal numbers and for similar pyramids of energy and biomass that Lindeman identified. Energy entered the community in primary production, and with something like 10% efficiency of transfer at subsequent feeding steps, was typically reduced to only 10^{-k} of its original amount by the k'th step. Having exhausted energy, the food chain could not continue, and standing crops of upper trophic levels were compatibly small, fitting consistently at the tops of the pyramids.

The Fig. 1 oyster model can be investigated in relation to this paradigm. The model contains 108 simple paths (Table 5), as computed by Ulanowicz (pers. comm.) using a modification of NETWRK (Ulanowicz 1982). Thirty-three of these are cycles (Table 3). In graph theory terms, a food chain is a simple acyclic path connecting adjacent compartments by sequential predator–prey relations, i "eats" j, in which energy is transferred from each j to each i in the sequence. Of the $108 - 33 = 75$ noncyclic simple paths in the oyster system, 20 are food chains: eight of length 1 ($1 \rightarrow 6$, $2 \rightarrow 3$, $2 \rightarrow 4$, $2 \rightarrow 5$, $3 \rightarrow 4$, $3 \rightarrow 5$, $4 \rightarrow 5$ and $5 \rightarrow 6$), seven of length 2 ($2 \rightarrow 3 \rightarrow 4$, $2 \rightarrow 3 \rightarrow 5$, $2 \rightarrow 4 \rightarrow 5$, $2 \rightarrow 5 \rightarrow 6$, $3 \rightarrow 4 \rightarrow 5$, $3 \rightarrow 5 \rightarrow 6$ and $4 \rightarrow 5 \rightarrow 6$), four of length 3 ($2 \rightarrow 3 \rightarrow 4 \rightarrow 5$, $2 \rightarrow 3 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 4 \rightarrow 5 \rightarrow 6$, and $3 \rightarrow 4 \rightarrow 5 \rightarrow 6$), and one of length 4 ($2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$). Of these food chains, only five are unique: $1 \rightarrow 6$, $2 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 3 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 4 \rightarrow 5 \rightarrow 6$ and $2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$. Note that there are two food sources, the bivalves themselves (compartment 1) and the detrital pool (compartment 2).

Table 6 shows the Lindeman "progressive efficiencies" and their consequences for the five model food chains. The upper section shows efficiencies for the arc components of these food chains computed from the flow data in Fig. 1. The lower portion computes efficiencies for the five chains as products of those of the appropriate arc components in each chain. These efficiencies give the amounts of energy remaining at the tops of each chain per unit of energy entering at the base. These are: $1 \rightarrow 6$, 1.2%; $2 \rightarrow 5 \rightarrow 6$, 0.2%; $2 \rightarrow 3 \rightarrow 5 \rightarrow 6$, 0.4%; $2 \rightarrow 4 \rightarrow 5 \rightarrow 6$, 0.2%; and $2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$, 0.03%. The last two columns show the reduction of energy in absolute terms from the bottom to the top of each chain. It is clear that this model conforms to the classical expectation. Energy is rapidly diminished as it passes along each food chain, so there is very little left at the top, and energy pyramids are produced.

A similar result is obtained by a path and flow analysis based on Tables 1a and b when storage is precluded from the system. Energy flow paths under such conditions can be represented as powers $A(0)^k$ of matrix $A(0) = A - I$, the Table 1a matrix with zero diagonal elements. With no flow contributions to storage, a corresponding non-dimensional flow matrix $P(0)$ can be constructed by zeroing the diagonal entries of P in Table 1b, and analyzing resultant higher order flows as $P(0)^k$. Table 7a illustrates the results for all food chains $1 \rightarrow \ldots \rightarrow 6$ and $2 \rightarrow \ldots \rightarrow 6$ of selected path lengths. Only 6.419×10^{-5} and 1.133×10^{-6} of each unit of energy introduced in compartments 1 and 2, respectively, ever reaches compartment 6. This is consistent with, although even more drastic than, the results obtained above by trophic level transfer efficiency analyses. Most of the energy introduced into the oyster reef model system without storage flushes rapidly through, being dissipated after only one or two trophic transfers. Efficiency of energy transfer is indicated to be exceedingly low in this case.

TABLE 5. Simple paths in the oyster reef model; cycles are identified by asterisks (*).

21 Paths from Node 1

1→2
1→2→3
1→2→3→4
1→2→3→4→5
1→2→3→4→5→6
1→2→3→5
1→2→3→5→6
1→2→4
1→2→4→5
1→2→4→5→6
1→2→5
1→2→5→6
1→6
1→6→2
1→6→2→3
1→6→2→3→4
1→6→2→3→4→5
1→6→2→3→5
1→6→2→4
1→6→2→4→5
1→6→2→5

21 Paths from Node 2

2→3
2→3→4
*2→3→4→2
2→3→4→5
*2→3→4→5→2
2→3→4→5→6
*2→3→4→5→6→2
2→3→5
*2→3→5→2
2→3→5→6
*2→3→5→6→2
2→4
*2→4→2
2→4→5
*2→4→5→2
2→4→5→6
*2→4→5→6→2
2→5
*2→5→2
2→5→6
*2→5→6→2

19 Paths from Node 3

3→4
3→4→2
*3→4→2→3
3→4→2→5
3→4→2→5→6
3→4→5
3→4→5→2
*3→4→5→2→3
3→4→5→6
3→4→5→6→2
*3→4→5→6→2→3
3→5
3→5→2
*3→5→2→3
3→5→2→4
3→5→6
3→5→6→2
*3→5→6→2→3
3→5→6→2→4

18 Paths from Node 4

4→2
4→2→3
*4→2→3→4
4→2→3→5
4→2→3→5→6
*4→2→4
4→2→5
4→2→5→6
4→5
4→5→2
4→5→2→3
*4→5→2→3→4
*4→5→2→4
4→5→6
4→5→6→2
4→5→6→2→3
*4→5→6→2→3→4
*4→5→6→2→4

17 Paths from Node 5

5→2
5→2→3
5→2→3→4
*5→2→3→4→5
*5→2→3→5
5→2→4
*5→2→4→5
*5→2→5
5→6
5→6→2
5→6→2→3
5→6→2→3→4
*5→6→2→3→4→5
*5→6→2→3→5
5→6→2→4
*5→6→2→4→5
*5→6→2→5

12 Paths from Node 6

6→2
6→2→3
6→2→3→4
6→2→3→4→5
*6→2→3→4→5→6
6→2→3→5
*6→2→3→5→6
6→2→4
6→2→4→5
*6→2→4→5→6
6→2→5
*6→2→5→6

How can the results of Tables 6 and 7a, corresponding to conventional trophodynamic concepts, be reconciled with the finding of the preceding section which showed enough energy storage in the system to support continued cycling and sustained storage for more than 500 days over path lengths exceeding $k = 2000$?

CONSEQUENCES OF STORAGE

The answer is simply that food chains are not the relevant units of energy transfer in ecosystems. Neither are food webs, because these also preclude nonfeeding transfers, such

as 1 → 2 (feces and pseudofeces deposition by the bivalves) in the oyster model. The relevant unit is the entire ecological network over which energy passes by any mechanism whatsoever. The network nodes must possess storage, however.

Patten (1984a) has shown in detail how the addition of storage to a flow network increases the number and diversity of paths and path types (with and without storages, cycles, nested storages and cycles, etc). This principle is illustrated in Table 7 for the two sets of food chains 1 → . . . → 6 and 2 → . . . → 6. For example, there are many more paths from compartment 1 to 6 than the single one accounted for in Table 6. When there is no storage, no length 2 paths exist from 1 to 6 (Table 7a), whereas there are two such paths (Table 7b) in the case with storage (1 → 1 → 6 and 1 → 6 → 6). Only one length 3 path (1 → 2 → 5 → 6) links 1 to 6 in the absence of storage, but four occur when storage is present (1 → 1 → 1 → 6, 1 → 1 → 6 → 6, 1 → 6 → 6 → 6 and 1 → 2 → 5 → 6). The total energy flow associated with length 3 paths increases from 8.390×10^{-10} when there is no storage to 7.081×10^{-7} when storage is accounted for. Length 10 paths increase from only 242 with associated flow 5.522×10^{-19} in the case without storage to 11032 with total flow 6.203×10^{-4} when storage is present. The same trends occur in the case of food chains leading from compartment 2 to 6. There is only one length 2 path; namely, 2 → 5 → 6; it contains no storage and is accounted for in the analysis of Table 6. Without storage there are only two length 3 paths (2 → 3 → 5 → 6 and 2 → 4 → 5 → 6) and an associated energy flow of 7.081×10^{-7}; with storage, however, there are three additional paths (2 → 2 → 5 → 6, 2 → 5 → 5 → 6 and 2 → 5 → 6 → 6), and altogether the five convey a total flow of 1.963×10^{-6}. There are 353 length 10 paths with associated energy 3.464×10^{-18} in the nonstorage case, and 16169 length 10 paths transferring 4.250×10^{-5} units of energy when storage is present.

In general, the following conclusions are demonstrated: (1) many more paths exist between any two compartments in a network with storage than in the same network without storage; (2) in consequence, there is much more energy flow over paths of a given length k with storage than without; (3) this is sufficient to prevent rapid flushing of energy through a system in only a few steps of transition time, and thus to hold energy within the system for the relatively long times associated with storage and transfer paths of great lengths. In conventional efficiency analysis of the oyster model (Table 6), paths of lengths

TABLE 6. Lindeman efficiencies of food chain energy transfer in the oyster reef model: energy flow to each feeder as a fraction of the input energy to its food compartment in each unique food chain in the model. Food chain efficiencies are products of compartment arc efficiencies.

Food Chain Arc Efficiencies

1→6: $f_{61}/z_1 = 1.238-2$	3→4: $f_{43}/f_{32} = 1.476-1$
2→3: $f_{32}/(f_{21} + f_{24} + f_{25} + f_{26}) = 3.670-1$	3→5: $f_{53}/f_{32} = 1.476-1$
2→4: $f_{42}/(f_{21} + f_{24} + f_{25} + f_{26}) = 3.267-1$	4→5: $f_{54}/(f_{42} + f_{43}) = 7.793-2$
3→5: $f_{52}/(f_{21} + f_{24} + f_{25} + f_{26}) = 2.889-2$	5→6: $f_{65}/(f_{52} + f_{53} + f_{54}) = 6.857-2$

Food Chains

Chains	Efficiencies		Energy Reduction (kcal m^{-2} d^{-1}) Base	Top
1→6	1.238−2		4.147+1	5.134−1
2→5→6	1.981−3		2.227+1	4.411−2
2→3→5→6	3.714−3	7.730−3	2.227+1	8.269−2
2→4→5→6	1.746−3		2.227+1	3.888−2
2→3→4→5→6	2.894−4		2.227+1	6.444−3

TABLE 7. Numbers of paths and quantities of nondimensional energy flows from compartments 1 and 2 to 6 for selected path lengths in the oyster reef model. Total flows to 6 over all paths of all lengths generated by one unit of energy in each initial compartment are listed at the bottom of each table. (a) The case without storage; (b) with storage.

	1→ . . . (all paths) . . . →6		2→ . . . (all paths) . . . →6	
Path lengths, k	Numbers of paths	Nondimensional energy flows	Numbers of paths	Nondimensional energy flows
		(a) *Nonstorage Case*		
1	1	6.419−5	0	0
2	0	0	1	4.251−7
3	1	8.390−10	2	7.081−7
10	242	5.522−19	353	3.464−18
50	4.604+15	5.465−81	6.748+15	7.897−80
100	1.836+32	3.429−158	2.691+32	4.955−157
500	—	0	—	0
2000	—	0	—	0
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
∞	∞	0	∞	0
Cumulative over all paths of all lengths		6.419−5		1.133−6
		(b) *Storage Case*		
1	1	6.419−5	0	0
2	2	1.279−4	1	4.251−7
3	4	1.911−4	5	1.963−6
10	11032	6.203−4	16169	4.250−5
50	9.194+23	2.693−3	1.347+24	9.444−4
100	7.320+48	4.574−3	1.073+49	2.334−3
500	—	6.650−3	—	4.461−3
2000	—	3.134−4	—	2.131−4
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
∞	∞	0	∞	0
Cumulative over all paths of all lengths		6.678		4.335

>4 are not considered because they would be compound (i.e., have repeated compartments), and this does not conform to the concept of a food chain (although, ironically, it would be consistent with the food cycles with which Lindeman's paper begins, but which in fact he never analyzed). At higher path lengths, say $k \geqq 10$ in the oyster model, huge numbers of energy transfer paths exist (Table 7), and these are completely omitted from Lindeman's analyses.

NETWORK AMPLIFICATION

As shown above, long paths even in small systems do carry energy, provided the

systems have storage. Table 7b indicates how much energy is transferred over paths of selected lengths in the oyster model and also gives the total nondimensional energy transferred over all paths of all lengths from compartments 1 and 2 to 6. For each unit of energy supplied to compartments 1 and 2 initially, 6.678 and 4.335 units, respectively, flow integrally to compartment 6 over all paths of all lengths in the system. The oyster model network thus has the property of "amplifying" original potential energy over time, and the Lindeman analysis, as well as that of Table 7a, is therefore seen to underestimate completely the true energy flow situation in ecosystem models.

This does not mean that more than one unit (the initial energy input) is ultimately respired, but it does mean that more than one unit of potential energy in the form of chemical bonds is stored and flowed during the time interval it takes for full dissipation of the initial unit to occur. Such network amplification of original energy is achieved by time integration.

Let the time dissipation of a unit of total system throughflow be described by

$$\frac{dT(k)}{dk} = -\lambda T(k), \quad T(0) = 1, \quad \lambda > 0,$$

whose solution is

$$T(k) = T(0)e^{-\lambda k} = e^{-\lambda k}.$$

The integral of $T(k)$ as $k \to \infty$ is

$$T = \int_0^\infty e^{-\lambda k}\, dk = -\lambda^{-1} e^{-\lambda k}\Big|_0^\infty$$

$$= -\lambda^{-1}(e^{-\lambda\infty} - 1) = \lambda^{-1}.$$

This time-integrated total system throughflow is greater for smaller values of the dissipation coefficient, and amplification results in the range $0 < \lambda < 1$. Therefore, greater integral throughflow is realized by slowing the dissipation rate.

In networks, this slowing is achieved through storage. Storage reduces directly the rate of energy loss and vastly increases the number and diversity of compound paths around which energy can cycle. This expands the ability of the energy to have nondissipative (in addition to dissipative) ecological effects as it circulates (Patten 1984a). Table 8 gives the Q matrix for the Fig. 1 model. The entries in each column denote time-integrated steady-state energy flows propagated to the row compartments i based on one unit of energy initially available in each column compartment j. The amplification achieved by this system is seen to be substantial, and it may be concluded that its various food chains are, therefore, not energy limited.

Food Chain and Trophic Level Limitations

If food chains are not energy limited, then why aren't they longer? Why are there so few trophic levels in ecosystems? Pimm (1982, p. 110 ff.) points out that size and design constraints exist on organism physiology and morphology which prevent the evolution of creatures that could cope with, e.g., lions, wolves, crocodiles or sharks as prey, even if the latter were abundant. With such fierce occupants of trophic level 5 or 6, an imaginative design process might yield feasible predators for yet several more levels, but not for the several hundred or thousand additional steps that energy transfer considerations alone would not prelude. The availability of exploitable potential energy at all trophic levels is met by parasites and their predators, as well as by detritivores and detritus food chains.

Widespread omnivory and an abundance of uncommon species produce complex, cross linked, many nodal networks from which chains can only be tediously extracted in the field. The process of abstracting these networks into models generates the appearance of short food chains.

The first step occurs in model aggregation. True complexities of real ecosystems cannot be represented isomorphically. Trophic models are produced by lumping species that occupy parallel positions in an energy flow network into relatively few compartments: producers, herbivores, detritivores, omnivores, carnivores, parasites, hyperparasites, etc. These trophic groups are then the nodes of relatively small order models.

The second step toward short food chains derives from the definition (implicit in empirical research) of food chains as simple, not compound, paths in networks. No matter how interconnected a low order, aggregated model may be in terms of energy flow (e.g., Fig. 1), the length of simple paths is necessarily small. In the oyster model the maximum simple path is the complete path of length 5: $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$.

The third step in establishing the shortness of food chains is provided by their strict definition as sequences of feeding relations, followed by unspoken restrictions on what constitutes "feeding." For example, the transfer $1 \rightarrow 2$ in the oyster system is mediated by two technically nonfeeding processes, food sorting to produce the pseudofeces component of detritus, and the repackaging of energy and nutrients for microbial consumption termed egestion (Dame and Patten 1981). Thus, the complete path is interrupted at this point, leaving only $2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ (length 4) as the longest food chain in the model. This sequence goes on for a very long time in terms of energy flow, $2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 2 \rightarrow \ldots$, but the registers that count food chain lengths are reset to zero each time the nonfeeding transfer $6 \rightarrow 2$ occurs. In these latter transfers, old energy is converted to new food, and so repackaged proceeds again along the chain. The energy cycles, but the food which the modeler, not the microbes, recognizes as such does not. Thus do food chains become, not energy limited, but food limited to short lengths.

Hypothesis 2. The principle that food chains are short and trophic levels few as consequences of inefficient energy transfer in ecosystems is false. These properties arise as artifacts of modeling, and from the failure to identify and account for all paths and associated energy flows from lower to upper trophic levels.

Direct vs. Indirect Interactions

Analogy with Economics

The network amplification property derived in the preceding section is reminiscent of attributes of economic systems. A unit of capital is invested against the various dissipative forces tending to erode it. If the investment is successful, the capital grows with time, and may be spent ("dissipated") or reinvested for continued growth. Energy entering the ecosystem represents initial cpaital, one unit to each compartment in the foregoing analyses. This energy immediately begins to be lost in respiration and other entropy generating processes, but what remains is "invested" in the network organization as storage. This storage slows the dissipation process, and also augments the network (more paths, cycles and storages in more diverse configurations) conferring the amplification property. By the time all initial energy is degraded, more than the original one unit of benefit or value has been realized by the system. The maximization of such value may be suggested as another optimality principle of ecosystems, like maximum power output (Odum and Pinkerton 1955), exergy (Jorgensen and Mejer 1979) or ascendency (Ulanowicz 1980), etc. Its measure is given directly by the transitive closure matrix (e.g., Table 8).

TABLE 8. Transitive closure matrix $Q = \sum_{k=0}^{\infty} P^k = (I - P)^{-1}$ for the oyster reef model.

From / To	1	2	3	4	5	6
1	1.929+2	0	0	0	0	0
2	9.646+1	2.494+2	4.968+1	1.401+2	1.977+2	1.187+2
3	2.327−1	6.017−1	1.301	3.380−1	4.769−1	2.863−1
4	2.327	6.017	2.882	1.476+1	4.769	2.863
5	1.570	4.059	4.937	4.302	2.915+1	1.931
6	6.678	4.335	5.273	4.594	3.113+1	4.060+2

Energy as Causality

The availability of long paths, even in small models like Fig. 1, suggests that indirect effects may be more significant than the direct interactions, such as competition and predation, that are of central concern in conventional ecology. Competition and predation are widely considered the two most important agents of adaptation to evolutionary fitness. The topic of indirect effects has been considered elsewhere using several different models (Patten 1982a, b, c, 1984c; Patten et al. 1982) of conservative flows, such as energy in the present case. These flows served as surrogates for causality by the following rationale.

Causality in real ecosystems is complex and mediated by many factors. A simple energy flow model cannot, and is not intended to, represent complex causation. However, if the latter could be quantified, and the quantities arrayed in a matrix as can readily be done for conservative substances, then properties revealed by analysis of the conservative flow matrices would be analogous to those associated with causal propagations. They are network properties rather than properties of what is passing over the networks. By such induction, the language of causation will be used in this section, even though it is the flow of energy that is investigated. What is revealed about energy relations will be taken as analogous to what would be learned by causal analysis were it possible to perform the latter.

Direct vs. Indirect Paths

A path between two compartments is *direct* if it involves only those compartments without intermediary exchange, and *indirect* otherwise. For example, $j \to j \to \ldots \to j \to i \to i \to \ldots \to i$ is direct, whereas $j \to \ldots \to k \to \ldots \to i$, $k \neq i$, j is indirect. A path such as $j \to j \to \ldots \to j \to i \to j \to j \to \ldots \to i$ may also be considered indirect, as influence passes to an intermediary compartment (in this case i) and then back to the originating compartment before passing finally to the terminal compartment. A cycle, $j \to \ldots \to k \to \ldots \to k \to \ldots \to i$, or storage, $j \to \ldots \to k \to k \to \ldots \to k \to \ldots \to i$, may appear within a path, or a nested cycle may appear within a cycle. The nesting of cycles and of cycle-storage combinations may be very complex over long paths. This is what was meant previously in referring to path diversification.

In this section the examination of causal paths in terms of direct vs. indirect effects will be limited to those associated with the food chains $1 \to \ldots \to 6$ and $2 \to \ldots \to 6$ considered earlier (Tables 6 and 7 and related text). These are but two examples of 28 ($6 \times 6 - 6$) nondiagonal interactions (Table 8) in the Fig. 1 system.

Table 7 shows the numbers of paths of selected lengths from compartments 1 and 2 to 6. As in the case of cycles (Table 2a), these paths grow combinatorially in number with increasing path length. The single length path from 1 to 6 ($1 \rightarrow 6$) is direct. There is no direct arc from 2 to 6. Two length 2 paths from 1 to 6 ($1 \rightarrow 1 \rightarrow 6$, $1 \rightarrow 6 \rightarrow 6$) are both direct by the above definition. The length 2 path from 2 to 6 ($2 \rightarrow 5 \rightarrow 6$) is indirect. Four length 3 paths occur from 1 to 6 ($1 \rightarrow 1 \rightarrow 1 \rightarrow 6$, $1 \rightarrow 1 \rightarrow 6 \rightarrow 6$, $1 \rightarrow 2 \rightarrow 5 \rightarrow 6$ and $1 \rightarrow 6 \rightarrow 6 \rightarrow 6$), all direct except the third. Five length 3 paths link 2 to 6 ($2 \rightarrow 2 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 3 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 4 \rightarrow 5 \rightarrow 6$, $2 \rightarrow 5 \rightarrow 5 \rightarrow 6$ and $2 \rightarrow 5 \rightarrow 6 \rightarrow 6$), all indirect. At path lengths ≥ 10 in the table, the numbers of paths are too large to identify specifically. There are, however, at path length k only k *direct* paths:

(1) $j \rightarrow j \rightarrow \ldots$ ($k - 2$ terms) $\rightarrow j \rightarrow i$

(2) $j \rightarrow j \rightarrow \ldots$ ($k - 3$ terms) $\rightarrow j \rightarrow i \rightarrow i$

$\vdots$

($k - 1$) $j \rightarrow j \rightarrow i \rightarrow i \rightarrow \ldots$ ($k - 3$ terms) $\rightarrow i$

(k) $j \rightarrow i \rightarrow i \rightarrow \ldots$ ($k - 2$ terms) $\rightarrow i$.

Therefore, $a_{ij}^{(k)} - k$ of the total $a_{ij}^{(k)}$ paths are indirect, and as k gets large, the number of indirect paths vastly exceeds the count of direct paths. With increasing path length $a_{ij}^{(k)}$ grows combinatorially, while k increases only linearly. This is the basis for the growth in importance of indirect effects as path lengths increase.

Direct vs. Indirect Effects

For the Markovian system $\hat{P}$ based on conventional throughflow, there will never be any path sequences of the form $j \rightarrow j \rightarrow \ldots \rightarrow j \rightarrow i \rightarrow i \rightarrow \ldots \rightarrow i$, because $\hat{p}_{jj} = \hat{p}_{ii} = 0$. Therefore, direct effects $\hat{d}_{ij}$ are defined by the simple direct interaction $\hat{p}_{ij}$, $\hat{d}_{ij} = \hat{p}_{ij}$, and indirect effects $\hat{i}_{ij}$ are the difference of this from total effects, $\hat{i}_{ij} = \hat{q}_{ij} - \hat{p}_{ij}$. This formulation has been applied previously to several different ecological models (Patten 1982a, b, c; Patten et al. 1982) based on inclusive throughflow before recognizing the distinction between conventional and inclusive throughflow. As a result indirect effects always appeared to greatly exceed the direct ones. Wiegert and Kozlowski (1984) criticized these results for containing pure storage sequences commingled with true indirect paths, and Higashi (1984) subsequently reformulated direct and indirect effects for the inclusive throughflow-based model.

As observed above, in the inclusive throughflow case there are k pure storage paths of each length k which must be discriminated from the $a_{ij}^{(k)} - k$ indirect paths. Direct, time delayed effects (storages) associated with the k length k paths estimated by $\sum_{m=1}^{k} p_{jj}^{k-m} p_{ij} p_{ii}^{m-1}$. Therefore, indirect effects of k'th order are $p_{ij}^{(k)} - \sum_{m=1}^{k} p_{jj}^{k-m} p_{ij} p_{ii}^{m-1}$. Total direct effects, d_{ij}, of all orders in the system, including initial conditions, are $d_{ij} = \sum_{k=0}^{\infty} \sum_{m=1}^{k} p_{jj}^{k-m} p_{ij} p_{ii}^{m-1}$. Higashi (1984) formulates these as $d_{ij} = p_{ij}/(1 - p_{ii})(1 - p_{jj})$, with the result that indirect effects are $i_{ij} = q_{ij} - p_{ij}/(1 - p_{ii})(1 - p_{jj})$. If one needs only the ratio of indirect to direct effects, or vice versa, Higashi (1984) has shown that these are the same regardless of whether the conventional or inclusive throughflow model is used. That is, $\hat{p}_{ij} = p_{ij}/(1 - p_{jj})$, and $\hat{q}_{ij} = q_{ij}(1 - p_{ii})$. Then, substituting these expressions into those for d_{ij} and i_{ij} above, $d_{ij} = \hat{p}_{ij}/(1 - p_{ii})$ and $i_{ij} = (\hat{q}_{ij} - \hat{p}_{ij})/(1 - p_{ii})$, and

TABLE 9. Matrix of indirect/direct effects ratios for the oyster reef model.

	Compartments					
From	1	2	3	4	5	6
To						
1	∞	0/0	0/0	0/0	0/0	0/0
2	0.410	∞	∞	0.560	0.488	0.388
3	∞	0.388	∞	∞	∞	∞
4	∞	0.619	0.716	∞	∞	∞
5	∞	4.419*	0.290	11.283*	∞	∞
6	0.299	∞	∞	∞	0.124	∞

*Indirect > direct.

$$\frac{i_{ij}}{d_{ij}} = \frac{(\hat{q}_{ij} - \hat{p}_{ij})/(1 - p_{ii})}{\hat{p}_{ij}/(1 - p_{ii})}, \text{ and } \frac{\hat{q}_{ij} - \hat{p}_{ij}}{\hat{p}_{ij}} = \frac{\hat{i}_{ij}}{\hat{d}_{ij}}.$$

Table 9 lists a matrix of these indirect/direct effects ratios for the oyster reef model. There are five indeterminate entries (0/0) in the table, indicating no direct or indirect influences. Of the remaining 31 elements, 19 (61.3%) show indirect/direct ratios of ∞. These are cases where indirect effects are exerted even though there are no corresponding direct interactions. Six (33.0%) of these indirect effects are stronger than the largest of the 18 (Table 1b) direct interactions in the model, and many others are comparable in magnitude to other direct causes. Thus, indirect influences may couple many non-interacting compartments as strongly as others are joined directly. In two cases (Table 9, asterisks) of compartments which interact directly, the strengths of these interactions are exeeded by those of the accompanying indirect influences conferred by membership in the system. Finally, based on data not shown, total indirect effects propagated globally in this model are 3.3 times greater than total direct effects. In no models examined to date has this characteristic failed to appear. These considerations indicate that indirect effects are clearly significant in comparison to direct effects in the Fig. 1 system.

Real ecosystems are of much higher order (many more compartments) than the 6th order oyster reef model. Patten et al. (1982) have shown that the combinatorics of both paths and effects are expressed sooner at smaller path lengths in such larger systems. Therefore, although indirect effects in the oyster models do not overwhelm direct ones, they are very significant, and scale considerations suggest this may be true a-forteriori in large, complex, real ecosystems. If so, then long term indirect effects may well dominate the direct interactions, such as competition and predation, traditionally considered important in evolutionary adaptation, niche differentiation, and other classical processes.

Hypothesis 3. The principle that direct interactions are the exclusive mediators of organism adaptations for fitness, and thus community organization, appears untenable in view of the apparent strength of indirect influences in ecosystems. The paradigm of biological evolution must be reexamined to take account of the contribution of indirect effects in natural selection.

Conclusion

Three generally held principles of contemporary ecology have been questioned through a systems-analytic examination of a small model of energy flow in an oyster reef community. The main conclusions have been stated as hypotheses pending more definitive investigations.

Hypotheses 1: Energy, like matter, cycles in ecosystems.

Hypothesis 2: Food chains are short and trophic levels few due to perspective and modeling conventions, not because of thermodynamic limitations.

Hypothesis 3: Direct effects may be less important in ecosystems than indirect interactions, compromising their utility in organizing biotic communities or mediating evolution.

The systems view of nature implicit in these results is more consistent with a physical or even mathematical, rather than a biological, conception of reality. The biosphere is organized into a syncytial network of energy-matter storages and flows, behaving and evolving according to strict laws that govern the organization of systems everywhere. The anthropomorphization of nature in the form of some current ecological excesses, such as strategies for adaptation, resource partitioning and other unneeded constructs diverts attention from the more basic principles involved in the causal (determinate and non-anticipatory) generation of reality.

Wangersky and Wangersky (1980, 1981, 1983), working in biological oceanography, have apparently unearthed one of these principles and have been kind in their assessment of more traditional explanations. Of the latter, they suggest they may be sufficient but unnecessary:

> Our model of benthic populations requires very few assumptions; our creatures eat, metabolize, reproduce and die. No assumptions are made concerning species interactions, except that those individuals nearest the food get to eat, and those few things they do, they all do alike. Any discussion of the effects of niche breadth or specialization is therefore irrelevant, and the only selection pressure working differentially on the species is that of the chance occurrence of food. These experiments with our artificial animals suggest that competition and niche specialization, as they are commonly understood, are not necessary for the establishment of mosaic-like aggregates in an apparently featureless environment. Yet at moderate food levels our species co-exist, and in patterns which strongly suggest those for which ecologists have commonly invoked environmental mosaics. We would not suggest that such mosaics do not exist, or that specializations for small environmental differences cannot produce such patterns. We do suggest that even in the absence of these factors, if food arrives randomly in time and space the pattern of local aggregation and apparent species separation will appear. It is not possible to distinguish specialization for small differences in niches from the effects of random food supply on the basis of either spatial distributions or counts of species and individuals. In mathematical terms, factors other than the food distribution may be sufficient but not necessary (Wangersky and Wangersky 1981, p. 784.)

From a random distribution of resources, the "manna effect" of these authors can produce local aggregations and stable separations into energy-matter storages. Direct interactions between these create networks of speciated and diversified paths of great numbers and great lengths (Patten 1984a on the path consequences of storage). These networks render unimportant the original direct interactions for any purpose other than to generate them, and once generated, the networks confer holistic properties resident in the mathematical laws of organization. These properties, in turn, shape the unitary development or coevolution of the entire ecosystem.

Such a systems paradigm of the natural world seems emergent in many areas of science in the present period. It can be expected to produce many revisions of how the world is viewed, including the observations about ecosystem energy cycling, lengths of food chains, and relative significance of direct vs. indirect effects made here. As the topical concerns of biological oceanography inevitably must intersect the ecosystem as a level of

organization, it seems prudent that this field should seek and support the strong introduction of systems thinking and systems analysis methodology into its future programs and activities.

Acknowledgements

Jine-i Kou authored software and performed the computations for A^A, P^A, Q and related data. Robert E. Ulanowicz computed Table 3 using NETWRK (Ulanowicz 1982). Masahiko Higashi clarified many of the fine points about direct vs. indirect effects. This research was supported by Grants BSR-8114823 and BSR-8215587 from the U.S. National Science Foundation.

References

DAME, R. F., AND B. C. PATTEN. 1981. Analysis of energy flows in an intertidal oyster reef. Mar. Ecol. Progr. Ser. 5: 115–124.

ELTON, C. 1971. Animal ecology. Macmillan, New York, NY.

FINN, J. T. 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol. 56: 363–380.

HANNON, B. 1973. The structure of ecosystems. J. Theor. Biol. 41: 535–546.

HIGASHI, M. 1984. Explicit expressions for direct and indirect effects in systems. Am. Nat. (Submitted)

HIGASHI, M., AND B. C. PATTEN. 1984. Role of storage in compartmental ecosystems. Ecol. Model. (Submitted)

JORGENSEN, S. E., AND H. MEJER. 1979. A holistic approach to ecological modelling. Ecol. Model. 7: 169–189.

KEMENY, J. G., AND J. L. SNELL. 1960. Finite Markov chains. Van Nostrand–Reinhold, Princeton, NJ.

LEONTIEF, W. W. 1966. Input–output economics. Oxford Univ. Press, London.

LINDEMAN, P. L. 1942. The trophic-dynamic aspect of ecology. Ecology 23: 399–418.

MATIS, J. H., AND B. C. PATTEN. 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Bull. Int. Stat. Inst. 48: 527–565.

ODUM, H. T., AND R. C. PINKERTON. 1955. Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. Am. Sci. 43: 331–343.

PATTEN, B. C. 1978. Systems approach to the concept of environment. Ohio J. Sci. 78. 206–222.

1982a. On the quantitative dominance of indirect effects in ecosystems, p. 27–37. *In* W. K. Lauenroth, G. V. Skogerboe, and M. Flug [ed.] Analysis of ecological systems: State-of-the-art in ecological modelling. Elsevier, Amsterdam.

1982b. Environs: relativistic elementary particles for ecology. Am. Nat. 119: 179–219.

1982c. Indirect causality in ecosystems: its significance for environmental protection. *In* W. T. Mason, and S. Iker, [ed.] Research on fish and wildlife habitat. U.S. Environmental Protection Agency, EPA-600/9-82-022, Washington, D.C.

1984a. Energy cycling in the ecosystem. Ecol. Model. (In press)

1984b. Further developments toward a theory of the quantitative dominance of indirect effects in ecosystems. Verh. Gesellschaft f. Ökologie, Bd. XI. (In press)

PATTEN, B. C., AND M. HIGASHI. 1984. Modified cycling index for ecological applications. Ecol. Model. 25: 69–83.

PATTEN, B. C., M. C. BARBER, AND T. H. RICHARDSON. 1982. Path analysis of a reservoir ecosystem model. Can. Water Res. J. 7: 252–282.

PATTEN, B. C., R. W. BOSSERMAN, J. T. FINN, AND W. G. CALE. 1976. Propagation of cause in ecosystems, vol. 4, p. 457–579. *In* B. C. Patten [ed.] Systems analysis and simulation in ecology. Academic Press, New York, NY.

PIMM, S. L. 1982. Food webs. Chapman and Hall, New York, NY.

ROBERTS, F. S. 1976. Discrete mathematical models. Prentice-Hall, Englewood Cliffs, NJ.

ULANOWICZ, R. E. 1980. An hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

1982. NETWRK: a package of computer algorithms to analyze ecological flow networks. Univ. of Maryland, Center for Environmental and Estuarine Studies, Chesapeake Biological Laboratory, Solomons, Maryland. UMCEES. Ref. No. 82-7 CBL.

WANGERSKY, P. J., AND C. P. WANGERSKY. 1980. The manna effect: a model of phytoplankton patchiness in a regenerative system. Int. Rev. Gesamten Hydrobiol. 65: 681–690.

1981. The manna effect: the structure of benthic populations. Int. Rev. Gesamten Hydrobiol. 66: 777–786.

1983. The manna effect: paradox of the plankton. Int. Rev. Gesamten Hydrobiol. 68: 327–338.

WIEGERT, R. G., AND J. KOZLOWSKI. 1984. Indirect causality in ecosystems. Am. Nat. 124: 293–298.

Flow analysis of Materials in the Marine Euphotic Zone

M. J. R. FASHAM

Institute of Oceanographic Sciences, (NERC), Brook Road, Wormley, Godalming, Surrey, GU8 5UB U.K.

1. Introduction

One of the themes running through a recent review of mathematical models in biological oceanography (Platt et al. 1981) was the importance of having at least as much information on the fluxes in an ecosystem as on the biomass. This review also stressed the need to place more emphasis on holistic as against reductionist models, and a number of candidate techniques for such holistic modelling were reviewed, of which the most relevant to the study of flows was input−output analysis (Hannon 1973; Patten et al. 1976; Hannon, this volume; Patten, this volume). In fact, in the marine field there have already been a few holistic approaches to the flow of energy or carbon through the food web (Petipa et al. 1970; Steele 1965, 1974; Mills 1979). However, none of these studies has quantitatively investigated the effect of material recycling, for which the mathematical technique of flow analyses or cycle analysis are a prerequisite. Furthermore, although the role of microplankton, both heterotrophic and autotrophic, in the system was often discussed, there was some uncertainty as to how to incorporate the microplankton into the models.

In 1974 Pomeroy suggested that new information on the role of microorganisms in the ocean was changing the traditional concept of a food web as being dominated by the phytoplankton−herbivore link, and that these new ideas had important implications both for the ultimate fate of primary production and for element recycling. Since then there has been an increasing amount of research on the ecology of nanoplankton, bacteria and protozoa, and a number of reviews have attempted to synthesize this information (Williams 1981a, 1984; Sorokin 1981; Joint and Morris 1982; Azam et al. 1983; Ducklow 1983).

In a parallel development there have been a number of attempts to view the marine ecosystem in terms of particle size distributions (Sheldon et al. 1972, 1977; Platt and Denman 1977; Steele and Frost 1977). The main idea underlying these theories is that the majority of zooplankton tend to eat particles that are less than a certain fraction of their own length but do not discriminate between particles of different taxonomic type. An extreme form of this concept was postulated by Isaacs (1972, 1973), who suggested that food webs in the ocean are completely "unstructured", that is, that most animals feed on whatever food, living or detrital, they are capable of catching or filtering. Fasham (1981) has shown how this concept can be formulated in terms of compartment models.

In this paper an attempt will be made to unite some of these new ideas on the structure of the marine food web with the concept of material recycling by means of an analysis of the carbon and nitrogen flows in an idealised marine food web.

2. A Model of the Pelagic Food Web

The model is intended to represent a steady-state food web in the euphotic zone of the ocean bounded below by either a permanent or seasonal nutricline. The model will thus apply to temperate and subtropical seas in the summer and to tropical seas throughout the whole year. At this stage no attempt will be made to define the model for a particular

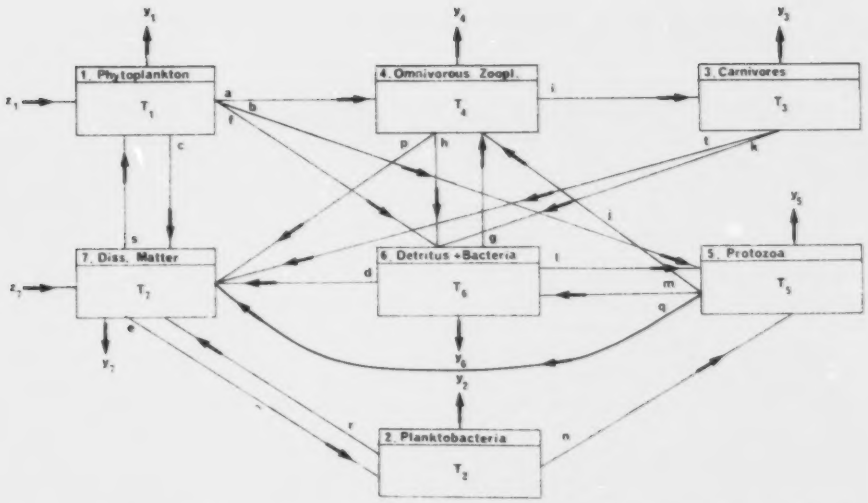


FIG. 1. The model of the marine euphotic zone ecosystem showing the inter-compartment flows, system inputs (z_1, z_7) and outputs ($y_1, y_2, \ldots y_7$). The parameters a, b, etc. represent the proportions of the total output from a compartment flowing along a given pathway, and the parameters $T_1, T_2, \ldots T_7$ are the throughputs of the compartments.

marine ecosystem type, i.e., neritic, oligotrophic, upwelling etc., although most of the data for defining the parameters and testing the results comes from neritic areas such as the North Sea.

The model consists of seven compartments (Fig. 1) which represent the following components of the ecosystem: (1) Phototrophic phytoplankton, (2) Planktobacteria, (3) Carnivores, (4) Omnivorous metazoic, zooplankton, (5) Protozoa, (6) Detritus plus epibacteria, and (7) Dissolved material. These categories will now be defined in greater detail.

2.1 Phototrophic Phytoplankton

This compartment consists of phototrophic prokaryotic and eukaryotic phytoplankton, which, by using light energy to fix dissolved inorganic carbon, initiate the whole production cycle. To maintain growth these cells also need to take up dissolved inorganic nitrogen and other nutrients. Recent results (Watt 1971; Platt et al. 1983; Joint and Pomroy 1983) have shown that between 30 and 60% of the total photosynthetic activity may be due to cells less than 1 μm in diameter (picoplankton) consisting mainly of blue–green algae. In this model it is assumed that the picoplankton are too small to be ingested by most metazoic zooplankton but will instead form part of the food supply of the protozoa (Revalante and Gilmartin 1983). It is further assumed that some proportion of the total primary production may pass directly into the detrital component (Walsh 1983; Billett et al. 1983) or be exudated as dissolved organic material (DOM) (Fogg 1983).

2.2 Planktobacteria

The planktobacteria are those bacteria capable of existing as free planktonic cells and not requiring a solid surface to grow on (Sieburth 1979). These bacteria grow by absorbing

dissolved organic carbon (DOC) through their cell membranes, but they also require nutrients, for which they may have to compete with the phytoplankton (Azam et al. 1983; Ducklow 1983). In the model it is assumed that the planktobacteria are preyed on only by the protozoa (Azam et al. 1983; Fenchel 1984), although some zooplankton such as *Oikopleura* are known to assimilate bacteria (King et al. 1980).

2.3 Carnivores

This category consists of animals that are capable of some active discrimination in their selection of food and are assumed to be obligate carnivores. Examples are the large crustacea, chaetognaths, siphonophores, and fish, covering the size range from a few mm to tens of meters. Because of their relatively large size, it is assumed that they cannot assimilate protozoa and are dependent on the omnivorous zooplankton for their food. As with most animals, their food supply is converted into a number of metabolic products which, for carbon flow, are: (i) respiration of CO_2, (ii) egestion of fecal pellets plus, for crustacea, the shedding of moults, (iii) excretion of DOC and (iv) growth plus reproduction.

2.4 Metazoic Omnivorous Zooplankton

This category consists of those animal groups, such as copepods, ostracods and euphausiids, that are assumed to be fairly omnivorous in their diet. The size of these animals ranges from 100 μm to a few cm and are, therefore, considered to be incapable of ingesting the pico-phytoplankton or planktobacteria. Their diet will consist of medium to large phytoplankton, protozoa and detritus.

2.5 Protozoa

This category consists mainly of flagellates and ciliates that are capable of ingesting planktobacteria and the smallest phytoplankton. It is also assumed that part of their diet will consist of small detrital particles (or the bacteria attached to these particles [Fenchel 1970]). Azam et al. (1983) and Ducklow (1983) have suggested that the main predation on bacteria is by flagellates, which are in turn preyed on by ciliates; but this added complication will not be included in the model.

2.6 Detritus Plus Epibacteria

This category consists of the non-living particulate organic matter, such as fecal pellets, discarded moults and dead organisms together with their attendant epibacteria. These bacteria, which originate mainly from the guts of the egesting animals (Sieburth 1979), can grow only on a particulate substrate, so they can be regarded structurally as one "organism" (Sorokin 1981). For many zooplankton it is probably the attached micro-organisms that constitute the food value of detritus (Fenchel 1970; Sorokin 1981). It is assumed that the activity of the epibacteria results in the loss of DOC and nutrients to the dissolved material compartment. Other loss terms from this compartment are the fall-out of detrital material from the euphotic zone and the respiration of CO_2 by the bacteria.

2.7 Dissolved Matter

This category is comprised of the dissolved fraction of the element being considered. In the case of carbon this is restricted to utilisable dissolved organic carbon (UDOC),

which may be only a small percentage of the total DOC. It is this UDOC which is utilised as a substrate by the planktobacteria (Azam and Ammerman 1984). It is assumed that a certain proportion of the organic carbon excreted by the animals may not be utilisable by bacteria, and this fraction is considered to be a loss to the refractory pool of DOC.

For nitrogen, the dissolved fraction will consist of both inorganic (NO_3, NO_2, and NH_3) and organic (urea, amino acids) forms. The inorganic nitrogen can be utilised directly by the phytoplankton, while the organic forms will be utilised by the planktobacteria, which will recycle part of it to the dissolved fraction as inorganic nitrogen. In the tropical and subtropical ocean throughout the year, and in the temperate ocean throughout the summer, there is a strong nutricline below the deep chlorophyll maximum. The process of vertical diffusion acting on this nutrient gradient will produce an upwards flux of inorganic nitrogen into the photic zone. This is the only major source of "new" (as against "regenerated") nutrients (Eppley 1981) into the photic zone, and this addition is modelled by a constant input to the dissolved nitrogen compartment.

3. The derivation of the Flow Equations

The mathematical basis of input–output flow analysis has been discussed in other papers in this symposium (Hannon, Patten, this volume), and only a brief summary will be given using the notation of Patten et al. (1976). Let the flow from compartments j to i be ϕ_{ij} and the input and output from compartment j be z_j and y_j, respectively. The throughflow of the kth compartment is defined as the sum of either the inflows,

$$(1) \quad T_k = \sum_{j=1}^{n} \phi_{kj} + z_k, \qquad k = 1, \ldots n,$$

or of outflows,

$$(2) \quad T'_k = \sum_{i=1}^{n} \phi_{ik} + y_k, \qquad k = 1, \ldots n,$$

where n is the total number of compartments. In a steady-state situation $T_k = T'_k$. The present interest in analysing the model of the euphotic zone described above is to determine the fate of the inputs, which are either the primary carbon assimilation in the case of the carbon cycle or the upwards flux of new nutrient in the case of the nitrogen cycle. This is achieved by what Patten et al. (1976) term a genon outflow analysis. We assume that the direct flow, ϕ_{kj}, can be expressed as a fraction of the throughflow in compartment j;

$$(3) \quad \phi_{kj} = q_{kj} T_j \qquad k = 0, 1, \ldots n,$$

where q_{kj} is the proportion of throughflow of T_j destined for compartment k and $\phi_{oj} \equiv y_j$. The critical assumption in this formulation is that the proportions q_{kj} are independent of the magnitude of T_j.

Substituting for ϕ_{kj} in equation (1) gives,

$$(4) \quad T_k = \sum_{j=1}^{n} q_{kj} T_j + z_k.$$

Let $\boldsymbol{T}$ and $\boldsymbol{Z}$ be column vectors of throughflows and inflows, respectively, and $\boldsymbol{Q}$ be a $n \times n$ fractional outflow matrix, then equation (4) can be written as

$$(5) \quad \boldsymbol{T} = \boldsymbol{Q}\boldsymbol{T} + \boldsymbol{Z},$$

or

$$(6)\quad \boldsymbol{T} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}\boldsymbol{Z}.$$

The matrix $\boldsymbol{S} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}$ is called the output structure matrix (Ulanowicz 1984), and the element s_{ij} represents the throughflow in compartment i generated by one unit of inflow into compartment j.

One of the features of interest in the model is the amount of recycling taking place in the system. For a compartment k, Finn (1978) defined a "return cycling efficiency" given by,

$$(7)\quad R_k = (s_{kk} - 1)/s_{kk}.$$

This quantity measures the fraction of the throughput of compartment k that eventually returns to that compartment.

Let us now apply these concepts to the model shown in Fig. 1. The compartments have all been assigned a number which will identify their row or column position in the matrices $\boldsymbol{Q}$, $\boldsymbol{S}$, and $\boldsymbol{T}$. The reason for the choice of numbers will be given later. In order to simplify writing the equations, the direct flows will be represented by the primed lower case letters $a', \ldots, n', p', \ldots t'$, (for example $\phi_{41} = a'$), and the outputs by $y_1', \ldots, y_7'$. The matrix $\boldsymbol{Q}$ contains the fractional outflows, and these will be represented by the same lower case letters but without the prime. Referring to Fig. 1 shows, that, for example,

$$a = a'/(a' + b' + c' + f' + y_1'),$$

and

$$y_1 = y_1'/(a' + b' + c' + f' + y_1').$$

Using the symbolism it is a simple matter to write down the matrix $(\boldsymbol{I} - \boldsymbol{Q})$ thus,

$$(8)\quad \boldsymbol{I} - \boldsymbol{Q} = \left[\begin{array}{ccccc:cc} 1 & 0 & 0 & 0 & 0 & 0 & -s \\ 0 & 1 & 0 & 0 & 0 & 0 & -e \\ 0 & 0 & 1 & -i & 0 & 0 & 0 \\ -a & 0 & 0 & 1 & -j & -g & 0 \\ -b & -n & 0 & 0 & 1 & -t & 0 \\ \hdashline -f & 0 & -k & -h & -m & 1 & 0 \\ -c & -r & -t & -p & -q & -d & 1 \end{array}\right]$$

It is now necessary to calculate the inverse of this last matrix, and in all previous applications of flow analysis this has been done numerically by assuming values for the fractional outputs and inverting the matrix on a computer. However, bearing in mind that nearly half of the entries in the $\boldsymbol{I} - \boldsymbol{Q}$ matrix are zero, one may invert this matrix analytically by using a few mathematical devices.

The first device is the technique of inverting a matrix by partitioning. Let a matrix $\boldsymbol{A}$ of order n and its inverse $\boldsymbol{B}$ be partitioned into four submatrices

$$\left[\begin{array}{c:c} \underset{(u \times u)}{\boldsymbol{A}_{11}} & \underset{(u \times v)}{\boldsymbol{A}_{12}} \\ \hdashline \underset{(v \times u)}{\boldsymbol{A}_{21}} & \underset{(v \times v)}{\boldsymbol{A}_{22}} \end{array}\right], \quad \left[\begin{array}{c:c} \underset{(u \times u)}{\boldsymbol{B}_{11}} & \underset{(u \times v)}{\boldsymbol{B}_{12}} \\ \hdashline \underset{(v \times u)}{\boldsymbol{B}_{21}} & \underset{(v \times v)}{\boldsymbol{B}_{22}} \end{array}\right],$$

where $n = u + v$. Then, provided $\boldsymbol{A}_{11}^{-1}$ is nonsingular, the submatrices of $\boldsymbol{B}$ are given by,

$$\boldsymbol{B}_{11} = \boldsymbol{A}_{11}^{-1} + (\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{22})\ \boldsymbol{\Sigma}^{-1}(\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}),$$

$$\boldsymbol{B}_{21} = -\boldsymbol{\Sigma}^{-1}(\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}),$$

$$\boldsymbol{B}_{21} = -(\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})\ \boldsymbol{\Sigma}^{-1},$$

$$\boldsymbol{B}_{22} = \boldsymbol{\Sigma}^{-1},$$

where $\boldsymbol{\Sigma} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}(\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})$.

Applying this stratagem to our problem, if the matrix $\boldsymbol{I} - \boldsymbol{Q}$ is partitioned as indicated by the dashed lines in equation (8), then instead of having to determine the inverse of a 7×7 matrix, it is necessary only to calculate the inverse of a 5×5 matrix, $\boldsymbol{A}_{11}$, and a 2×2 matrix, $\boldsymbol{\Sigma}$ — the subsequent manipulations being simply matrix multiplication. The inverses of $\boldsymbol{A}_{11}$ and $\boldsymbol{\Sigma}$ can be determined using the relationship:

$$\boldsymbol{A}^{-1} = \frac{\text{adj}\,\boldsymbol{A}}{|\boldsymbol{A}|},$$

where adj $\boldsymbol{A}$ is the adjoint of $\boldsymbol{A}$ and $|\boldsymbol{A}|$ its determinant. This is straightforward for the 2×2 matrix $\boldsymbol{\Sigma}$, while for $\boldsymbol{A}_{11}$ the derminant and adjoint can be found using the well-known expansion of determinants in terms of minors and cofactors. This in turn is made simpler if the matrix $\boldsymbol{A}_{11}$ contains a large number of zero elements, and it was for this reason that the Detritus and Dissolved Matter compartments were numbered 6 and 7, so that the large number of non-zero row entries (due to the fact that all the material cycling goes through these compartments) would not occur in the matrix $\boldsymbol{A}_{11}$. It also simplifies the algebraic manipulations if any rows of $\boldsymbol{A}_{11}$ for which all the off-diagonal elements are zero (i.e. Phytoplankton and Planktobacteria) are at the top of the matrix.

By using these devices it is possible, after some lengthy but straightforward algebraic manipulations, to obtain formulae for the elements of matrix $\boldsymbol{S}$ in terms of the parameters $a, b, \ldots$ etc., and these are given in the appendix. These formulae may seem rather complicated, but a simple computer program is all that is required to evaluate them. The throughflow in compartment i caused by a unit input into compartment j is given by element s_{ij} so that it is a straightforward matter to calculate the intercompartmental flows, outputs, and recycling efficiency indices produced by this input.

4. The Carbon Flow Model

4.1 Estimation of the Parameters

The goal of this flow analysis is to estimate the proportion of primary production, envisaged as the input z_1 of carbon to the phototrophic phytoplankton compartment, that flows through the various parts of the model ecosystem. To do this it is necessary first to determine values for the proportional output parameters, $a, b, \ldots$ etc. for each compartment. When this task was attempted, it soon became apparent that suitable data are not available for some compartments, while the existing literature gives conflicting estimates for others. The approach here is to select what appears to be the most recent estimates of the parameters and then to investigate the sensitivity of the results so obtained to changes in these estimated values. As at the moment there is no certainty that holistic models of this sort will be successful, this approach seems reasonable. Furthermore, even if it transpires that the models give realistic predictions, the parameter values would almost certainly be different in different types of marine ecosystems; values that may be suitable

for the Sargasso Sea may not be applicable to the North Sea. No attempt has been made to investigate this aspect of the model; but, as already mentioned, some of the data for testing the validity of the model comes from neritic seas with commercial fisheries, so the selection of parameter values was biased towards this type of ecosystem. The selection of parameter values for each compartment will now be discussed in more detail.

4.1.1 Phototrophic Phytoplankton

The carbon flow in the marine ecosystem is initiated by the photosynthetic fixation of dissolved CO_2 into the phytoplankton. It is generally agreed that photosynthesis in the ocean is never limited by the supply of CO_2, and it is for this reason that a dissolved inorganic carbon component was not included in the model. Part of the primary production is immediately lost through the respiration of CO_2 by the phytoplankton, and it is the remaining net production that is utilised by the rest of the marine food web. We will therefore define the input, z_1, to be the net primary production so that the output y_1 will be zero.

In a recent study in the Celtic Sea, Joint and Pomroy (1983) found that up to 75% of the primary production could be attributed to phytoplankton less than 5 μm. In the Sargasso Sea, Platt et al. (1983) found that between 50 and 60% of the production occurred in organisms less than 1 μm in size. The initial assumption in the model was that the proportion a and b were equal, and the sensitivity of the results to this assumption was also tested.

It is now generally agreed that some proportion of the primary production is lost from the cells by the exudation of DOC (Joint and Morris 1982; Fogg 1983). Experimental values for this proportion range from 1 to 50% (Parsons et al. 1977; Joint and Morris 1982), but an average value for healthy cells appears to be 15% and this value will be used here.

Finally, it is necessary to determine the proportion of the primary production ending up in phytoplankton cells that die and become part of the detritus without being egested by animals. There is some evidence that during the spring bloom in temperate seas this proportion may be quite high (Billett et al. 1983; Walsh 1983; Smetacek 1984). However, this model is intended to focus on the steady-state summer period, so an arbitrary value of 10% will be assumed for this loss. In summary, the final values for the phytoplankton parameters are $a = b = 0.375$, $c = 0.15$, $f = 0.10$.

4.1.2 Planktobacteria

The parameters to be estimated for the planktobacteria are the proportions of output going to growth (n) and excretion (r). Values between 40 and 90% have been reported for the growth efficiency of planktobacteria (Joint and Morris 1982; Azam et al. 1983). In the model a conservative value of 60% will be assumed, and the sensitivity of the results to this parameter will be tested. Azam et al. (1983) suggested that a negligible amount of carbon is excreted by planktobacteria, so the remaining 40% of assimilated carbon was assumed to be respired.

4.1.3 Omnivorous Metazoic Zooplankton

The zooplankton parameters are the growth efficiency (i), the proportion of food egested (h = 1-assimilation efficiency) and the proportion excreted (p) — the balance to be made up by respiration.

Raymont (1983) has reviewed the copious experimental data on assimilation efficiency. Values ranged from 60 to 95%, with the more recent experimental values at the higher end

of this range. Dagg (1976) determined a detailed carbon budget for the carnivorous amphipod *Calliopus laeviusculus*. He found that as much as 30% of the intake of prey might be lost by messy feeding. The proportion of organic carbon going to other sources (at 15°C) were: metabolism 22%, egestion 7%, moults 5%, growth and reproduction 29%, and leakage (i.e., excretion) 7%. In order to estimate the model parameters it is necessary to know the proportion of the 30% of prey lost by messy feeding that goes to DOC versus detritus. The experiments of Dagg (1974) and Copping and Lorenzen (1980) suggested that a large proportion goes to DOC. However, as Dagg (1974) points out, animals kept in small experimental enclosures can recapture large parts of the prey that are broken off during feeding. However, this recovery is less likely to occur in the ocean, so I have assumed arbitrarily that 10% of the messy feeding goes to detritus and 20% to DOC. If the leakage figure is added to the DOC proportion and the moults to the detritus, then this suggests that 27% of carbon input goes to DOC and 22% to detritus. In view of the similarity of these two figures and the various uncertainties, it will be assumed that the proportions going to DOC and detritus are equal; the actual value will be assigned below.

Experiments to estimate growth efficiency (i) have also yielded a wide range of values from 5% to 35% (Conover 1968). The value of 10% has often been assumed in the past to calculate the flow of energy in trophic pyramids (Slobodkin 1961), but Steele (1965, 1974), Cushing (1977), and Parsons et al. (1977) have argued that the efficiency of marine invertebrates may be as high as 20%. I will assume a compromise value of 15%.

It remains only to determine the proportion of carbon that is respired. Suschenya (1970) reviewed a number of experiments that gave a range for the percentage of energy expenditure on respiration of between 40 and 85% of assimilated energy. If parameter values of $i = 0.15$, and $h = p = 0.25$ were used, this would imply that the respiration y_4 was 0.35, or 47% of assimilated carbon. This figure is rather to the low end of the reported range, so the final chosen parameters were $i = 0.15$, and $h = p = 0.20$. This assumption results in 56% of assimilated carbon being expended on respiration.

4.1.4 Carnivores

The carnivorous animals are considered to be active searchers after food and must therefore expend more enegy on metabolism (Sharp 1984). Growth efficiencies of 10% are generally assumed for fish (Steele 1974; Mills 1979), and it will be assumed that the 5% decrease in growth efficiency (compared to the carnivourous zooplankton) is lost to respiration. The value chosen for growth efficiency implies that 30% of the input carbon goes to excretion and egestion, and it will again be assumed that the losses are divided equally between the two processes.

It is now necessary to decide how to close the model, i.e., to decide the fate of the 10% growth yield. In an unexploited ecosystem this growth yield would be removed by mortality with the dead animals becoming part of the detrital compartment. In the analyses of exploited ecosystems, however, it is often assumed that half the yield goes to natural mortality and half to fisheries (Steele 1974). The fishery yield is a loss to the system and must be added to the respiration, y_3. As the model results will be compared with data from the North Sea, I will assume that the model system is being exploited, giving parameter values for this compartment of $k = 0.25$ and $t = 0.20$.

4.1.5 Protozoa

The growth efficiency of protozoa was assumed to be 40% (Fenchel 1982), while the proportion of input going to respiration was assumed to be the same as for bacteria (40%).

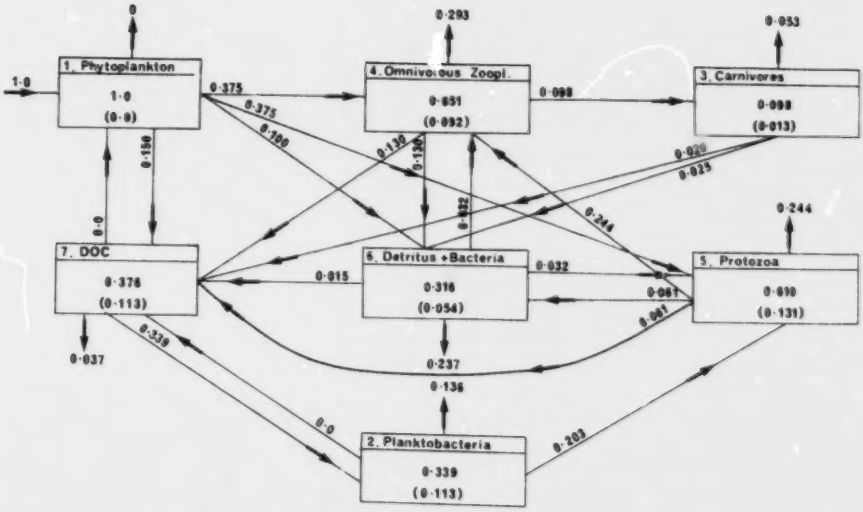


FIG. 2. The carbon flows produced by unit net primary production of the phytoplankton compartment. The unbracketed and bracketed figures inside each compartment box are the throughput and cycling efficiency index, respectively, for that compartment. The parameter values for this flow structure were: $a = b = 0.375$, $c = 0.15$, $d = 0.15$, $e = 0.90$, $f = 0.10$, $g = 0.10$, $h = 0.20$, $i = 0.15$, $j = 0.40$, $k = 0.25$, $\ell = 0.10$, $m = 0.10$, $n = 0.60$, $p = 0.20$, $q = 0.10$, $t = 0.20$ and $r = s = 0.0$.

The remaining 20% was divided up equally between excretion and egestion, giving the parameter values, $j = 0.40$ and $m = q = 0.10$.

4.1.6 Dissolved Organic Carbon (DOC)

It was assumed that no DOC is taken up directly by phytoplankton ($s = 0$) and that 10% of the input is lost to the refractory pool, giving a value of 0.9 for e. Although the 10% loss factor was an arbitrary guess, it is interesting to note that this assumption will eventually result in a 3.7% loss of primary production to the refractory pool (Fig. 2), which agrees well with the independent estimate (3%) made by Skopintsev (1971).

4.1.7 Detritus and Epibacteria

The parameters for the detrital compartment were the most difficult to determine. It is first necessary to estimate the proportion of carbon lost from the system (y_6), which is the sum of the detrital fall-out and the epibacterial respiration. Then the proportion reingested by animals and lost to DOC must be determined.

The proportion of primary production lost as fall-out from the euphotic zone will obviously vary with area and season. Thus, Deuser (1971) estimated that 10% was lost in the Black Sea, while in the tropics a figure of 4.4% has been estimated (Small et al. 1983). For the North Sea, Steele (1965, 1974) has variously assumed that 20 and 30% of the annual primary production reached the bottom to provide food for the demersal fish.

In order to develop some feel for this problem, some preliminary calculations were made to study the effect on y_6 of varying the parameters d, g, and ℓ; the other parameters were as previously defined. For the parameter values $g = \ell = 0.2$, $d = 0.1$ it was found that y_6 was 17% of primary production which, bearing in mind that this must account for bacterial respiration, is probably too low. On the other hand, if all three parameters were set to zero, then the loss rate was 30%. A compromise between these two parameter sets was assumed, and values $g = \ell = 0.1$, $d = 0.05$ were chosen giving a loss rate of 24%. If it is rather conservatively assumed that half this loss is detrital fall out, then 12% of primary production would be lost from the euphotic zone as fall-out. This is close to Deuser's figure but much lower than the values assumed by Steele.

This loss rate will also be affected by the choice of the egestion parameters, n, k, and m, as will be seen in the section on sensitivity analysis.

4.2 Results of the Standard Model

The parameter values estimated in the previous section can be used to derive values for the elements s_{i1}, $i = 1, \ldots 7$ of the output structure matrix using the equations given in the appendix. This will constitute the standard model. The values of the s_{i1} are the throughputs of carbon in the compartments 1 to 7 caused by a unit input of carbon into compartment 1. How this throughput is divided between the various outputs from a compartment is then calculated from the proportional output parameters a, b, etc., and finally the recycling index for each compartment is calculated using equation 7. The values of all these derived variables for the standard model are shown in Fig. 2.

The first point to notice is the importance of the microbial contribution to the total community respiration. Respirations by the planktobacteria and protozoa compartments account for 38% of the primary production, while the omnivorous zooplankton and carnivore compartments account for 34%. Williams (1981b) measured the size distribution of respiration in two mesocosms, and found that the respiration of organisms smaller than 10 μm could be up to 80% of the total for all organisms less than 330 μm in size. In the model a comparable figure would be the ratio of planktobacteria plus protozoa respiration to the respiration of the omnivorous zooplankton. This ratio for the standard model is 56%, which lies within the range of values observed by Williams (1981b).

Andrews and Williams (1971) and Fuhrman and Azam (1980) concluded from radioactive tracer experiments that between 20–30% or more of the primary production could be utilised by microorganisms. This flux is represented by the flow from DOC to planktobacteria, which in the standard model is slightly higher at 34% of net primary production. Duursma (1962) estimated that the total yearly production of DOC in the southern North Sea was 52 g C/m^2. If the total particulate primary production is estimated as 90 g C/m^2 (Steele 1974), then the DOC production is 58% of this value. The comparable figure from the model, the ratio of throughput in the DOC compartment (0.376) to the particulate primary production (0.85), is 44%.

These comparisons show that the model estimates of carbon flow through the microbial loop (from DOC through planktobacteria to protozoa) are broadly consistent with recent experimental data and give support to the many microbiologists who have been stressing the importance of this loop in recent years (Pomeroy 1974; Williams 1981a; Joint and Morris 1982; Azam et al. 1983). The results of the model, however, provide some information on two aspects of the microbial loop for which experimental data is not readily available, namely the source of the input of DOC and the importance of the microbial loop to the higher end of the food web.

It has long been recognised that input to the DOC pool could come from phytoplankton

exudation, animal excretion or breakdown of POC, but the relative importance of these sources was ambiguous (Steele 1974). Williams (1981a) suggested that 53% came from phytoplankton exudation (a high value of 0.3 was assumed for the parameter c), 29% from animal excretion and 18% from POC. Joint and Morris (1982) did not consider the breakdown of POC as a source of DOC and assumed that 82% of DOC came from animal excretion and only 18% from phytoplankton exudation. The results of the flow model give the proportions 56% from animal excretion, 40% from phytoplankton exudation and 4% from POC, which lie somewhere in between the conjectures of Williams (1981a) and Joint and Morris (1982). Obviously, these proportions will be affected by the choice of parameters, particularly the parameters a, p, q, and t. As an example the exudation parameter c was increased to 0.3, the figure assumed by Williams (1981a), and the parameters a, b, and f reduced to 0.325, 0.325, and 0.05, respectively. These changes resulted in an increased flow through the microbial loop, and the proportions of DOC coming from exudation, animal excretion and POC were 58, 39, and 3%, respectively. The percentage due to exudation is now similar to that calculated by Williams (1981a), although the percentage due to POC is still considerably lower than his value. The reason for this discrepancy is that Williams (1981a) did not allow for any loss of detritus out of the euphotic zone, nor for any recycling back to the animals. The effect of altering the excretion parameters will be studied in more detail in the next section, but it should already be clear that changes in these parameters will affect the relative proportions of DOC throughput that are contributed by plants and animals. However, the results from the standard model would support Joint and Morris (1982) in stressing the importance of animal excretion as the main source of DOC.

Most previous quantitative analyses of the role of microorganisms in the food chain have tended to concentrate on the problem of providing enough food for their various estimated production requirements (Joint and Morris 1982; Williams 1981a) rather than on the fate of these productions. In the present model it has been assumed that the production from the protozoan compartment is entirely consumed by the omnivorous zooplankton. If this assumption is correct, then the model predicts that as much as 37% of the throughput of this compartment is obtained from the protozoa. If true, this result is further evidence of the importance of the microbial loop to the marine food chain. One of the reasons that this proportion is so high is that high growth yields of 60 and 40% were assumed for planktobacteria and protozoa, respectively. If both these yields are reduced by 20%, then the carbon throughput of the omnivorous zooplankton is reduced by 22%, and the proportion coming from the protozoa is only 20% of the total. However, even in this case, it is clear that the contribution of the microbial loop to the higher food chains, and thus ultimately to fish yields, cannot be ignored.

Let us now consider the possible fish yield to man from this model food web. I previously assumed that 5% of the carnivore throughput (i.e., half of yield) could be exported from the system as yield to fisheries. The carnivore throughput is 0.098, which would give a fish yield of 0.5% of the net primary production or 0.6% of the particulate primary production. This may be compared with a yield of North Sea pelagic fish of 4 $kcal \cdot m^{-2} \cdot yr^{-1}$, which is 0.4% of the primary particulate production of 900 $kcal \cdot m^{-2} \cdot yr^{-1}$ (Steele 1974). Steele (1965, 1974) produced a simple energy flow model of the North Sea that could provide the observed pelagic fish yield (and the yield of demersal fish) from the observed primary production. It was necessary, however, to assume a growth yield of 19% for the herbivore population, and furthermore the energy requirements for bacteria and protozoa were not catered for. The model described here can produce the required fish yield by assuming only a 15% growth yield for the omnivorous zooplankton, and it is also capable of supporting a substantial production of bacteria and

protozoa (20 and 24% of the primary production, respectively). The main differences between the two models that produce this apparent increase in total system yield are:

(1) Steele's model did not consider the fate of excreted organic carbon or of fecal pellets remaining in the euphotic zone. A high proportion of this material can be recycled, either directly via animal detritivory or more indirectly via the microbial loop. The recycling efficiency index (Fig. 2) shows that 9% of the carbon entering the omnivore compartment had been recycled.

(2) Steele assumed that the animals at the first step in the food chain were totally herbivorous, whereas here it has been assumed that they are omnivores that will eat all particles (large plants, detritus and protozoa) of a certain size class.

(3) Steele assumed that half of the herbivorous production (i.e., omnivorous production in this model) would be eaten by invertebrate carnivores, which were in turn eaten by fish. This introduces a further trophic level (see also Mills 1979), and the effect of this on the model results can be roughly estimated, if it is assumed that the invertebrate carnivores have a 10% growth efficiency and that the effect of their waste products is minimal. The throughput of the carnivore compartment is then reduced from 0.098 to 0.054, giving a fish yield of 0.3% of the net particulate primary production, i.e., three-quarters of the value observed in the North Sea.

(4) Steele (1974) assumed that one third of the net primary production was lost from the euphotic zone as sinking detritus. The present model predicts a loss of 24% of primary production, which loss comprises both detrital sinking and epibacterial respiration. If the modelled detrital fall-out is again assumed to be half the total compartment loss, then this would be a third of Steele's figure.

These intercomparisons raise a number of issues which will need further study and evaluation of available experimental data. However, I think it may be concluded that the carbon flow analysis of the simple marine food web introduced in this paper demonstrates that it is possible to support both a substantial microbial population and a reasonable fish yield, as long as the recycling of carbon through the DOC and detrital compartments is taken into account.

4.3 Sensitivity Analysis

The carbon flow model shown in Fig. 1 has 16 independent parameters, and it would obviously be very time consuming to investigate the sensitivity of the throughputs to all of these parameters. Instead attention will be focussed on three specific variations of the standard model. These are:

(i) Altering the growth efficiency of bacteria and protozoa.
(ii) Altering the ratio of net/nano phytoplankton.
(iii) Altering the relative proportions of animal egestion to excretion.

4.3.1 The Effect of Altering the Growth Efficiency of Bacteria and Protozoa

One of the important features of the standard model was the high growth efficiencies assumed for planktobacteria and protozoa. Such effective growth results in the efficient recycling of DOC that would otherwise be lost from the system. These efficiencies, though based on good experimental evidence, are much higher than have been assumed in most previous work (Slobodkin 1961; Steele 1965, 1974; Cushing 1977). Thus, it is important to investigate the consequences of altering these growth efficiencies. This was done by jointly altering the values of the parameters n and j over the ranges 0.15−0.75 and 0.10−0.50, respectively, while keeping the other parameters constant. The changes in the

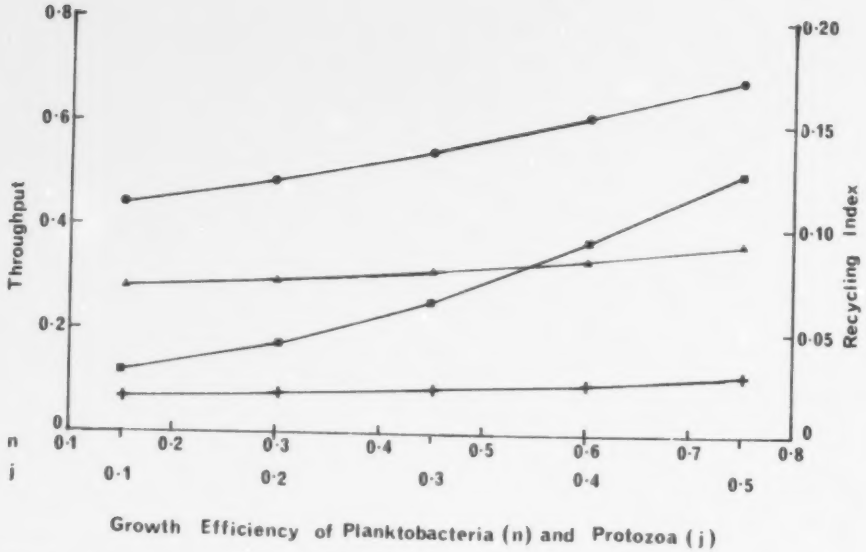


FIG. 3. The steady-state carbon throughput of the planktobacteria (▲), protozoa (●) and carnivore (+) compartments as a function of growth efficiency of planktobacteria (n) and protozoa (j). Also shown is the recycling efficiency index (■) for the omnivorous zooplankton compartment.

growth yield were assumed to be balanced by respiration. In Fig. 3 the values of the throughflows of the planktobacteria (T_2), protozoa (T_5) and carnivore (T_3) compartments, and the recycling index for the omnivorous zooplankton compartment have been plotted against the values of n and j. As the growth efficiency is increased, the throughput increases in all three compartments; but this increase is relatively greater for the carnivores (72%) than for the bacteria (32%). The reason for this asymmetry is the dramatic increase in recycling through the omnivore compartment from 3 to 12.5%, which again demonstrates the potential importance of the microbial loop to the marine food chain. However, a number of authors (Williams 1981a; Joint and Morris 1982; Azam et al. 1983; Ducklow 1983) have suggested that, if the bacteria have high growth efficiencies, then they are likely to be inefficient at recycling nutrients. This point will be discussed further in the section on the nitrogen flow model.

4.3.2 The Effect of Altering the Size Distribution of Phytoplankton

The parameters a and b represent the proportion of net primary production due to large (net plankton) and small (pico- and nano-plankton) phytoplankton, respectively. In the standard model these proportions were assumed to be 0.375 for both size classes. The effects on the throughputs T_3, T_4, and T_5 of changing these proportions from 100% large to 100% small phytoplankton production is shown in Fig. 4. The most obvious result is that, as production is switched to the small size class, there is a gradual reduction in the throughput of the omnivorous zooplankton and carnivore compartments and an increase in the throughput of the protozoa; the throughput of the planktobacteria is hardly altered.

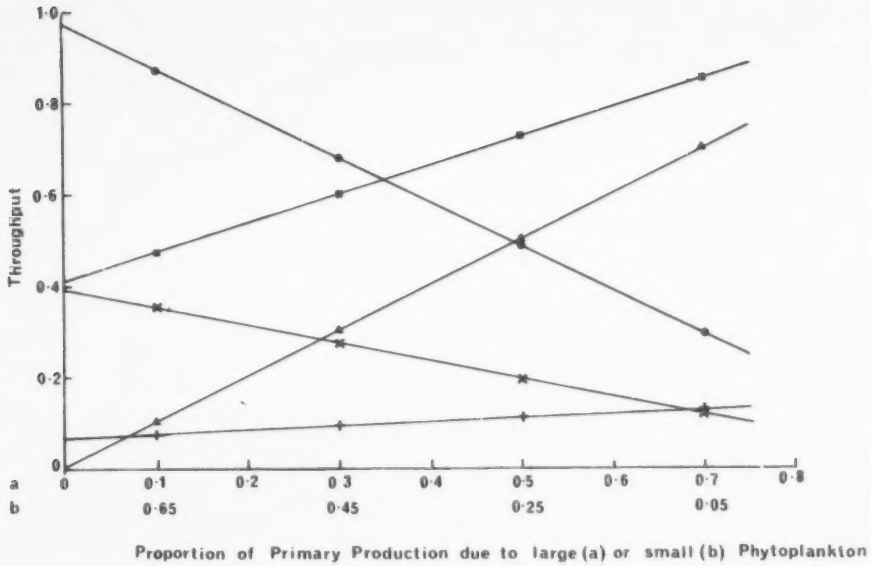


FIG. 4. The steady-state carbon throughput of the protozoa (●), omnivorous zooplankton (■) and carnivore (+) compartments as a function of the proportion of primary production due to large (=a) and small (=b) phytoplankton. Also shown are the input to the omnivorous zooplankton compartment from phytoplankton (▲) and protozoa (×).

These results are as expected; however it is worth noting that, although the throughput of the omnivore compartment is reduced by a half as a is decreased from 0.75 to 0.0, the input of carbon to this compartment from the protozoa increases to become the major food input to the omnivorous zooplankton. Thus, although one would expect an overall reduction in omnivorous zooplankton production (when the phytoplankton consists predominantly of the small size range), those groups of animals within the omnivores that were efficient at feeding on protozoa might actually increase in production.

As far as the yield of fish is concerned, the model predicts a lower yield (as a fraction of primary production) when the phytoplankton consist mainly of nano- and picoplankton. This result is interesting bearing in mind Steele's (1965) observation that, although the annual primary production in the Sargasso Sea may be as much as half that of the North Sea, the fish yield in the Sargasso Sea is considerably less. This might be partly explained by the fact that the production in the Sargasso Sea is mainly due to coccolithophores whereas larger diatoms predominate in more northerly waters (Watt 1971).

4.3.3 Alteration of the Relative Magnitude of Animal Egestion to Excretion

In the standard model the proportions of animal food intake going to excretion and egestion were assumed to be the same. The sensitivity of the model to this assumption was tested by altering the values of the parameters p, q, t and h, m, k so as to change this

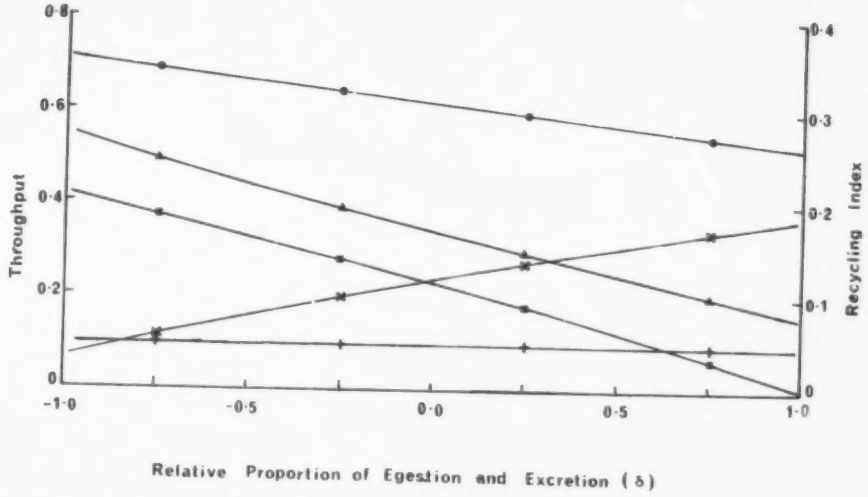


FIG. 5. The steady-state carbon throughput of the planktobacteria (▲), protozoa (●) and carnivore compartments (+) as a function of the parameter δ. δ = 1 represents 100% egestion, 0% excretion; and δ = −1 represents 100% excretion, 0% egestion (see text for details). Also shown are the system output (y_6) from the detritus and epibacteria compartment (×) and the cycling efficiency index for the DOC compartment (■).

proportion. The formula used for, say p and h, was

$$p = 0.2\,(1 - \delta),$$

$$h = 0.2\,(1 + \delta),$$

with δ varying between −1 and +1. The same formula was used for the pairs t, k and q, m, and the effect on the throughputs of the carnivore, planktobacteria and protozoa compartments is shown in Fig. 5. Switching from total excretion to total egestion produced a 72% reduction in planktobacteria throughput and a 27% reduction in protozoa throughput. However, there was rather surprisingly only a 6% reduction in carnivore throughput. The reason for this is that, with the standard values for the parameters associated with the detrital, DOC, planktobacteria and protozoa compartments, the carbon loss around the omnivorous zooplankton−DOC−planktobacteria−protozoa−omnivorous zooplankton loop was 21%, whereas around the omnivore−detritus−omnivore it was 25%. Thus, there is only a very small reduction in the recycling index for the omnivore compartment (0.092 to 0.073) on going from an excretion dominated to an egestion dominated situation. However, if different assumptions were made, say about the bacterial growth efficiency or detrital loss rate, then the outcome might be quite different. Nonetheless, it does show that quite major changes in the metabolism of the animals in the higher part of the food chain may be detectable only in the compartments normally considered to be at the bottom of the food chain, such as planktobacteria and protozoa.

Another obvious consequence of an increase in egestion is the increased loss of carbon from the euphotic zone as detrital fall-out. This loss y_6 ranged from 17% of primary production for 100% excretion to 38% for 100% egestion. This change demonstrates that

the ratio of excretion to egestion is a critical factor in determining the amount of primary production that will be available to support the deep ocean and benthic communities.

5. The Nitrogen Flow Model

The results from the carbon flow model were sufficiently encouraging to consider attempting to model the flow of nitrogen. Eppley and Peterson (1979) have pointed out that the nitrogen cycle is driven by inorganic nitrate diffusing up through the nutricline, giving rise to "new" primary production, in contrast to the "regenerated" primary production produced by nutrients recycled in the euphotic zone. This situation can be reproduced by driving the nitrogen flow model with an input into the dissolved nitrogen compartment.

The determination of the parameters for the nitrogen flow model presents some extra problems. In the carbon cycle model all the dissolved fraction was considered to be organic; the inorganic dissolved carbon was defined as being external to the model, as the supply of CO_2 is not generally considered to be limiting for oceanic photosynthesis. However, in the nitrogen cycle the dissolved fraction consists of inorganic nitrogen, such as NH_3, NO_2 and NO_3 ions, which can be taken up directly by the phytoplankton, and organic nitrogen, such as urea and amino acids, which may be taken up directly, but which will more likely be utilised by planktobacteria which may regenerate part of their organic uptake in inorganic form. Furthermore, a number of authors (Joint and Morris 1982; Azam et al. 1983) have suggested that the planktobacteria might compete with phytoplankton for inorganic nitrogen. All these processes have been parameterized in the model simply (probably too simply) by the choice of the three parameters e, s, and r.

The second problem arises when the nitrogen flows predicted by the model are compared with those of the carbon model. It has been pointed out succinctly by Joint (1983) that the C:N ratios produced by any model of C and N flows must lie within the range of observed experimental values. In fact, if the modelled flows for one element are taken as fixed, then this imposes a number of constraints on the flows of the other element. In this study the carbon flow model was considered to be based on more accurate experimental data; and, therefore, the nitrogen model was adjusted to give reasonable C:N ratios. It should be stressed that this adjustment was not done in any rigorous or optimal way, so the results should be regarded only as provisional. The reasons for the choice of the parameter values will now be given.

5.1 Estimation of the Parameters

5.1.1 Phytoplankton

The compounds exudated by phytoplankton are largely non-nitrogenous (Fogg 1983), and only a small proportion (5%) of the nitrogen throughput was assumed to be returned to the dissolved nitrogen compartment. The output of nitrogen produced by particulate primary production was assumed to be in the same proportions as for the carbon fractions, giving the parameter values $a = b = 0.419$, $f = 0.112$ and $c = 0.05$.

In order to calculate the C:N ratios in the model it is necessary to define the C:N ratio of the phytoplankton production, and the Redfield ratio (Goldman et al. 1979) 6.625 was used. The equations given in the appendix were used to calculate the throughputs in each compartment that are produced by a unit input into the Dissolved Nitrogen compartment.

If the throughput induced in the phytoplankton compartment is T_1, then if all the nitrogen flows are multiplied by the quantity $0.85/(6.625(0.95T_1))$, the resulting flows will be normalised by the Redfield ratio for phytoplankton production. Simple division will then give the C:N ratios from the carbon flows given in Fig. 2.

5.1.2 Planktobacteria and Dissolved Nitrogen

The growth efficiency of planktobacteria for nitrogen was assumed to be the same as that for carbon (0.6), and nitrogen not assimilated was assumed to be excreted. There do not appear to be any data with which to estimate the parameters e and s. It was found, however, that the C:N ratio of planktobacteria production was very sensitive to the parameter e. This parameter was, therefore, adjusted to give a reasonable C:N ratio of 5.3 (Joint 1983), and the remaining dissolved nitrogen flow (minus 10% to the refractory component) was assumed to go to the phytoplankton. The final values chosen were $e = 0.29$, $s = 0.61$.

5.1.3 Omnivorous Zooplankton, Carnivores, and Protozoa

In the initial attempts at deriving the nitrogen flow the same growth efficiency for omnivores was assumed as for carbon, and the remaining nitrogen output flow was divided equally between egestion and excretion — for which there is some experimental supporting evidence (Butler et al. 1970). However, it became apparent that this assumption produced C:N rations of ~3 for the fecal material, whereas recent experimental observations suggested values in the range 5−7.5 (Small et al. 1983). Accordingly, the fraction of output going to excretion was increased (while maintaining the same growth efficiency), until a value in the observed range was obtained. The final parameter values chosen were: $i = 0.15$; $p = 0.675$; $h = 0.175$; $j = 0.40$; $q = 0.525$; $m = 0.075$; $t = 0.675$; $k = 0.275$.

5.1.4 Detritus and Epibacteria

The values of g and ℓ for detritus were arbitrarily assumed to be the same as in the carbon flow model, and then the parameter d, representing the remineralisation of detritus by the epibacteria, was adjusted until the C:N ratio of the detrital fall-out was approximately 10 (Knauer and Martin 1981). In calculating this C:N ratio it was again assumed that 50% of the carbon detritus sank from the euphotic zone and 50% was respired by the epibacteria. The final parameter values chosen were $d = 0.375$ and $g = \ell = 0.2$.

5.2 Results

The nitrogen flows obtained using these parameters are shown in Fig. 6. The recycling efficiency index shows that 85% of the dissolved nitrogen throughput comes from recycling. This is within the range of values quoted by Eppley and Peterson (1979) for open ocean ecosystems, but is probably a little too high for a neritic ecosystem, such as the North Sea. The recycling efficiencies of the other compartments lie between 65% and 80%, except for detritus (53%) and carnivores (31%). It is of interest to note the fractions of the recycled dissolved nitrogen contributed from different sources. Azam et al. (1983) and Ducklow (1983) have suggested that, because the planktobacteria are efficient at converting DOC to growth, they are likely to be inefficient at recycling nutrients — assuming that the growth efficiency for carbon and nitrogen are similar. However, the

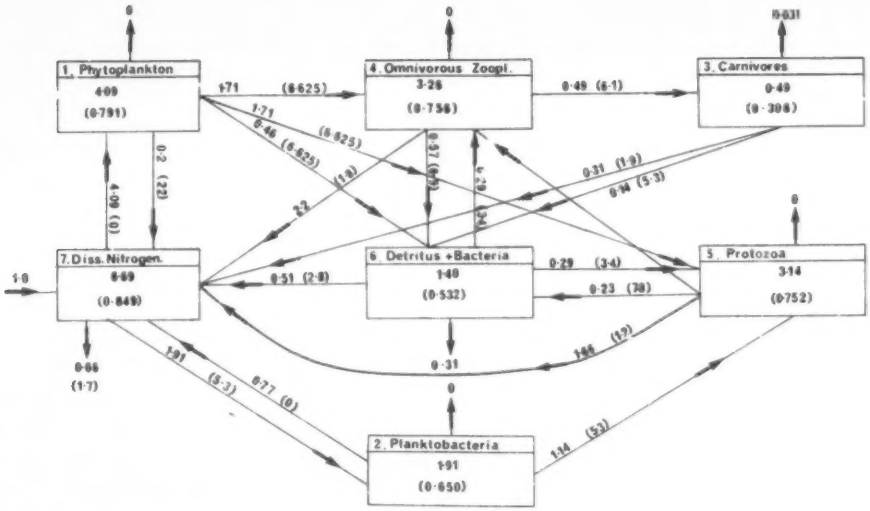


FIG. 6. The nitrogen flow produced by an input of 1.00 unit into the dissolved nitrogen compartment. The unbracketed and bracketed figures inside each compartment box are the throughput and cycling efficiency index, respectively, for that compartment. The figures in brackets after each intercompartment flows are the C:N ratios, and the flows have been normalised to give a C:N of 6.625 for the primary particulate production. The parameter values for this flow structure were: $a = b = 0.419$, $c = 0.050$, $d = 0.375$, $e = 0.290$, $f = 0.112$, $g = 0.200$, $h = 0.175$, $i = 0.150$, $j = 0.400$, $k = 0.275$, $\ell = 0.200$, $m = 0.075$, $n = 0.500$, $p = 0.675$, $q = 0.525$, $t = 0.675$, $r = 0.400$, and $s = 0.610$.

results of the nitrogen flow model show that the C:N ratio of the dissolved organic material taken up by the planktobacteria is low (2.4), so that if this uptake is to be used to synthesise cell material with a C:N ratio of 4−5, then it is obvious that the planktobacteria must either excrete a large amount of nitrogen, or alternatively take up less dissolved nitrogen. If the latter possibility is true, as was assumed in the model, then the planktobacteria may not be important remineralisers of nutrients. Thus, using the results given in Fig. 6, the planktobacteria produce 13% of the recycled nitrogen compared with 29% for protozoa, 39% for omnivorous zooplankton, 9% for epibacteria, and 5% for carnivores. These results, therefore, support Azam et al. (1983), who suggested that if the planktobacteria had a high growth efficiency, then the protozoa would make a significant contribution to the nutrient recycling. However, they also demonstrate that the contribution of zooplankton may be even more important than the protozoa. On the other hand, if the planktobacteria take a larger proportion of the dissolved nitrogen flow than has been assumed in the model, they would then need to excrete more nitrogen and would, therefore, assume more importance in nutrient recycling. Which of these possibilities is true may depend on the ratio of organic to inorganic dissolved nitrogen.

The C:N ratios for egested material and detrital fall-out were used to adjust the parameter values, so no further comment on these flows is needed. It is interesting to note that the C:N ratio decreases along the food chain phytoplankton → omnivorous zooplankton → carnivores, which is consistent with an increased protein content of fish compared to crustacea and plants.

6. Discussion

This attempt at applying the technique of flow analysis to the carbon and nitrogen flows of a simple model of the marine euphotic zone must perforce be regarded as a preliminary exercise. In cases where no information was available the choice of the parameters was made on arbitrary grounds, and in the remaining cases the values were selected after a somewhat personal perusal of the available literature. However, having started out on this journey with no certainty of arriving anywhere, I personally feel that the results are very encouraging.

The mathematical technique of flow analysis enables one to estimate correctly and unambiguously the inter-compartment flows, a nontrivial problem that can cause some confusion when the degree of recycling is large (see e.g. Joint (1983) and Newell and Field (1983)). In the case of the carbon cycle, the flow analysis demonstrates that it is possible to support both a reasonable fish yield and bacterial production, if the recycling of DOC is taken into account. The ability to deal with recycled flows is even more important for the nitrogen cycle, where up to 86% of the flow is recycled. However, if these flows are calculated correctly, they can be combined with the carbon flow data to calculate C:N ratios, which provides a check on the choice of parameter values.

Before the success (or otherwise) of this approach can be assessed, a number of problems will require further research. Probably the most critical of these is to provide some justification for lumping all carnivores or all protozoa into one compartment. If the food webs within these categories are unstructured, then the theoretical analyses of Isaacs (1972, 1973) and Fasham (1983) may provide a rationale, but some further theoretical work is needed here.

There is an obvious requirement to more thoroughly review the literature for estimates of the model parameters and for observations on flow rates that can be used to validate the model. The choice of compartments may need further consideration. For instance, the detritus plus epibacteria compartment might be more usefully split into separate units; similarly the dissolved nitrogen might be split into organic and inorganic components.

Another obvious step would be to consider what changes in parameter values would be required to model either different ecosystems (e.g., neritic, tropical, oligotrophic) or different seasons. This last possibility raises the necessity for a time dependent flow analysis (Patten et al. 1976). Also, the model could be extended to include both the water column beneath the euphotic zone and the benthos.

The question of model sensitivity has been only briefly touched on in this paper and obviously requires further study. It is a straightforward, if laborious, task to calculate the partial differentials $\partial T_i / \partial x$ (where x is any parameter) using the equations given in the appendix. This approach may prove very useful for, say, investigating the critical paths in nutrient cycling. A slightly more difficult problem concerns the development of a more rigorous or optimal method of constraining two flow models so as to give the correct element ratios in certain selected flows.

The technique of flow analysis does not provide any explicit estimation of the amount of biomass in the various compartments. By taking this viewpoint it avoids any necessity to investigate the biomass-specific details of the dynamics in and between the various compartments. However, in view of the copious data on biomass values it may prove useful to attempt to estimate the steady-state biomass of the individual compartments. There would seem to be three possible methods of doing this:

(1) Define the relationship between flows and biomass for each compartment and then solve for steady-state biomass using the standard techniques of compartmental dynamics. This is the approach used by Jones (1982) for the analysis of nutrient recycling.

(2) If independent estimates of turnover time for a compartment are available, biomass can then be estimated from the product of throughput and turnover time.

(3) If the flow rates for a particular compartment are accepted, the biomass might be calculated from a physiological model for that particular compartment using weight specific formulae for metabolic rates and a knowledge of the size distribution.

The technique of flow analysis as applied to a simplified compartmental model of a marine food web provides an holistic viewpoint that can help the development of unifying concepts. However, it remains to be seen through comparisons of model results and observations whether the assumptions of flow analysis and the aggregation of the multifarious species into a few compartments can be justified.

Acknowledgements

I would like to thank D. H. Cushing, J. H. Steele, and I. Joint for their comments on the first draft of this manuscript.

References

ANDREWS, P., AND P. J. LE B. WILLIAMS. 1971. Heterotrophic utilization of organic components in the sea. III. Measurement of the oxidation rates and concentrations of glucose and amino acids in water. J. Mar. Biol. Assoc. U.K. 51: 11–25.

AZAM, F., AND J. AMMERMAN. 1984. Cycling of organic matter by bacterioplankton in pelagic marine ecosystems: microenvironmental considerations, p. 345–360. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY.

AZAM, F., T. FENCHEL, J. G. FIELD, J. S. GRAY, L.-A. MEYER-REIL, AND F. THINGSTAD. 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Progr. Ser. 10: 257–263.

BILLETT, D. S. M., R. S. LAMPITT, A. L. RICE, AND R. F. C. MANTOURA. 1983. Seasonal sedimentation of phytoplankton to the deep-sea benthos. Nature 302: 520–522.

BUTLER, E. I., E. D. S. CORNER, AND S. M. MARSHALL. 1970. On the nutrition and metabolism of zooplankton. VII. Seasonal survey of nitrogen and phosphorus excretion of *Calanus* in the Clyde Sea area. J. Mar. Biol. Assoc. U.K. 50: 525–560.

CONOVER, R. J. 1968. Zooplankton — life in a nutritionally dilute environment. Am. Zool. 8: 107–118.

COPPING, A. E., AND C. J. LORENZEN. 1980. Carbon budget of a marine phytoplankton herbivore system with carbon-14 as a tracer. Limnol. Oceanogr. 25: 873–882.

CUSHING, D. H. 1975. Marine ecology and fisheries. Cambridge University Press, Cambridge. 278 p.

DAGG, M. J. 1974. Loss of prey body contents during feeding by an aquatic predator. Ecology 55: 903–906.

1976. Complete carbon and nitrogen budgets for the carnivorous amphipod, *Calliopius laeviusculus* (Krøyer). Int. Rev. Gesamten Hydrobiol. 61: 297–357.

DEUSER, W. G. 1971. Organic-carbon budget of the Black Sea. Deep-Sea Res. 18: 995–1004.

DUCKLOW, H. W. 1983. Production and fate of bacteria in the oceans. Bioscience 33: 494–501.

DUURSMA, E. K. 1962. The production of dissolved organic matter in the sea, as related to the primary gross production of organic matter. Neth. J. Sea Res. 2: 85–94.

EPPLEY, R. W. 1981. Autotrophic production of particulate matter, p. 343–363. *In* A. R. Longhurst [ed.] Analysis of marine ecosystems. Academic Press, London, New York.

EPPLEY, R. W., AND B. J. PETERSON. 1979. The flux of particulate matter to the deep ocean. Nature 282: 677–680.

FASHAM, M. J. R. 1981. Analytic food-web models, p. 54–65. *In* T. Platt, K. H. Mann, and R. E. Ulanowicz [ed.] Mathematical models in biological oceanography. The Unesco Press, Paris.

FENCHEL, T. 1970. Studies on the decomposition of organic detritus derived from the turtle grass *Thalassia testudinum*. Limnol. Oceanogr. 15: 14–20.

1982. Ecology of heterotrophic microflagellates. II. Bioenergetics and growth. Mar. Ecol. Prog. Ser. 8: 225–231.

1984. Suspended marine bacteria as a food source, p. 301–316. *In* M. J. R. Fasham [ed.] The flow of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY.

FINN, J. T. 1978. Cycling index: a general definition for cycling in compartment models, p. 138–165.

In D. C. Adriano and I. L. Brisbin [ed.] Environmental chemistry and cycling processes. ERDA.

Fogg, G. E. 1983. The ecological significance of extracellular products of phytoplankton photosynthesis. Bot. Mar. 26: 3–14.

Fuhrman, J. A., and F. Azam. 1980. Bacterioplankton secondary production estimates for coastal waters of British Columbia, Antarctica and California. Appl. Environ. Microbiol. 39: 1085–1095.

Goldman, J. C., J. M. McCarthy, and D. G. Peavey. 1979. Growth rate influence on chemical composition of phytoplankton in oceanic waters. Nature 279: 210–215.

Hannon, B. 1973. The structure of ecosystems. J. Theoret. Biol. 41: 535–546.

Isaacs, J. D. 1972. Unstructured marine food-webs and "pollutant analogues". Fish. Bull. 70: 1053–1059.

1973. Potential trophic biomasses and trace substance concentration in unstructured marine food-webs. Mar. Biol. 22: 97–104.

Joint, I. R. 1983. Bacterial contributions to kelp bed consumers: a comment on a paper by Newell and Field. Mar. Biol. Lett. 4: 245–248.

Joint, I. R., and R. J. Morris. 1982. The role of bacteria in the turnover of organic matter in the sea. Oceangr. Mar. Biol. Ann. Rev. 20: 65–118.

Joint, I. R., and A. J. Pomroy. 1983. Production of picoplankton and small nanoplankton in the Celtic Sea. Mar. Biol. 77: 19–27.

Jones, R. 1983. Nutrient recycling and ecosystem models. I.C.E.S. CM 1983/L: 32, 11 p.

King, K. R., J. T. Hollibaugh, and F. Azam. 1980. Predator–prey interactions between the larvacean *Oikopleura dioica* and bacterioplankton in enclosed water columns. Mar. Biol. 56: 49–57.

Knauer, G. A., and J. H. Martin. 1981. Primary production and carbon-nitrogen fluxes in the upper 1500m of the northeast Pacific. Limnol. Oceanogr. 26: 181–185.

Mills, E. L. 1980. The structure and dynamics of shelf and slope ecosystems off the north east coast of North America, p. 25–48. *In* K. R. Tenore and B. C. Coull [ed.] Marine benthic dynamics, University of South Carolina Press, SC.

Newell, R. C., and J. G. Field. 1983. Relative flux of carbon and nitrogen in a kelp-dominated system. Mar. Biol. Lett. 4: 249–257.

Parsons, T. R., M. Takahashi, and B. Hargrave. 1977. Biological oceanographic processes. Pergamon Press, Oxford, 332 p.

Patten, B. C., R. W. Bosserman, J. T. Finn, and W. G. Gale. 1976. Propagation of cause in ecosystems, p. 458–579. *In* B. C. Patten [ed.] Systems analysis and simulation in ecology, Vol. IV. Academic Press, New York, NY.

Petipa, T. S., E. V. Pavlova, and G. N. Mironov. 1970. The food web structure, utilization and transport of energy by trophic levels in the planktonic communities, p. 143–167. *In* J. H. Steele [ed.] Marine food chains. Oliver and Boyd, Edinburgh.

Platt, T., and K. Denman. 1977. Organisation in the pelagic ecosystem. Helgol. Wiss. Meeresunters. 30: 575–581.

Platt, T., K. H. Mann, and R. E. Ulanowicz [ed.] 1981. Mathematical models in biological oceanography. The Unesco Press, Paris. 157 p.

Platt, T., D. V. Subba-Rao, and B. Irwin. 1983. Photosynthesis of picoplankton in the oligotrophic ocean. Nature 301: 702–703.

Pomeroy, L. R. 1974. The ocean's food web, a changing paradigm. Bioscience 24: 499–504.

Raymont, J. E. G. 1983. Plankton and productivity in the oceans. Vol. 2 — Zooplankton. Pergamon Press, Oxford. 824 p.

Revelante, N., and M. Gilmartin. 1983. Microzooplankton distribution in the Northern Adriatic Sea with emphasis on the relative abundance of ciliated protozoans. Oceanol. Acta 6: 407–415.

Sharp, G. D. 1984. Ecological efficiency and activity metabolism, p. 459–474. *In* M. J. R. Fasham [ed.]. The flow of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY.

Sheldon, R. W., A. Prakash, and W. H. Sutcliffe. 1972. The size distribution of particles in the ocean. Limnol. Oceanogr. 17: 327–340.

Sheldon, R. W., W. H. Sutcliffe, and M. A. Paranjape. 1977. Structure of pelagic food chain and relationship between plankton and fish production. J. Fish Res. Board Can. 34: 2344–2353.

Sieburth, J. M. 1979. Sea microbes. Oxford University Press. New York, NY. 491 p.

Skopintsev, B. A. 1971. Recent advances in the study of organic matter in the oceans. Oceanology. 11: 775–789.

Slobodkin, L. B. 1961. Growth and regeneration of animal populations. Holt, Rinehart and Winston, New York, NY.

Small, L. F., S. W. Fowler, S. A. Moore, and J. Larosa. 1983. Dissolved and fecal pellet carbon and nitrogen release by zooplankton in tropical waters. Deep-Sea Res. 30: 1199–1220.

Smetacek, V. 1984. The supply of food to the benthos, p. 517–548. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY.

Sorokin, Y. I. 1981. Microheterotrophic organisms in marine ecosystems, p. 293–342. *In* A. R. Longhurst [ed.] Analysis of marine ecosystems. Academic Press, London, New York.

Steele, J. H. 1969. Some problems in the study of marine resources. Int. Comm. N.W. Atlant. Fish. Spec. Publ. No. 6, 463–476.

Steele, J. H. 1974. The structure of marine ecosystems. Blackwell Scientific Publications, Oxford. 128 p.

STEELE, J. H., AND B. W. FROST. 1977. The structure of plankton communities. Phil. Trans. R. Soc. 280: 485–524.

SUSCHENYA, L. M. 1970. Food rations, metabolism and growth of crustaceans, p. 127–141. *In* J. H. Steele [ed.] Marine food chains. Oliver and Boyd, Edinburgh.

ULANOWICZ, R. E. 1984. Community measures of marine food webs and their possible applications. p. 23–48. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York.

WALSH, J. J. 1983. Death in the sea: enigmatic phytoplankton losses. Progr. Oceanogr. 12: 1–86.

WATT, W. D. 1971. Measuring the primary production rates of individual phytoplankton species in natural mixed populations. Deep-Sea Res. 18: 329–339.

WILLIAMS, P. J. LE B. 1981a. Incorporation of microheterotrophic processes into the classical paradigm of the planktonic food web. Kiel. Meeresforsch. 5: 1–28.

1981b. Microbial contribution to overall marine plankton metabolism: direct measurements of respiration. Oceanol. Acta 4: 359–364.

1984. Bacterial production in the marine food chains: the emperor's new suit of clothes, p. 271–300. *In* M. J. R. Fasham [ed.] Flows of energy and materials in marine ecosystems: theory and practice. Plenum Press, New York, NY.

Appendix. Derivation of Flow Equations

The throughput matrix $\boldsymbol{T}$ is given by,

$$\boldsymbol{T} = [\boldsymbol{I} - \boldsymbol{Q}]^{-1}\boldsymbol{Z} = \boldsymbol{S}\boldsymbol{Z},$$

where the matrix $\boldsymbol{I} - \boldsymbol{Q}$ is

$$\left[\begin{array}{ccccc:cc} 1 & 0 & 0 & 0 & 0 & 0 & -s \\ 0 & 1 & 0 & 0 & 0 & 0 & -c \\ 0 & 0 & 1 & -i & 0 & 0 & 0 \\ -a & 0 & 0 & 1 & -j & -g & 0 \\ -b & -n & 0 & 0 & 1 & -\ell & 0 \\ \hdashline -f & 0 & -k & -h & -m & 1 & 0 \\ -c & -r & -t & -p & -q & -d & 1 \end{array}\right]$$

This matrix will be inverted by partitioning it into four submatrices, $\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$, $\boldsymbol{A}_{21}$, and $\boldsymbol{A}_{22}$, as indicated by the dotted lines. The inverse $\boldsymbol{S}$ will be partitioned into the submatrices $\boldsymbol{S}_{11}$, $\boldsymbol{S}_{12}$, $\boldsymbol{S}_{21}$, and $\boldsymbol{S}_{22}$. It is first necessary to calculate the inverse of $\boldsymbol{A}_{11}$ by calculating its adjoint and determinant. Then,

$$\boldsymbol{A}_{11}^{-1} = \frac{\operatorname{adj} \boldsymbol{A}_{11}}{|\boldsymbol{A}_{11}|} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ i(a + jb) & ijn & 1 & i & ij \\ a + jb & jn & 0 & 1 & j \\ b & n & 0 & 0 & 1 \end{bmatrix}$$

The next step is calculate the matrix,

$$\Sigma = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}(\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}) \equiv \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

The elements of this matrix are given by:

$$\sigma_{11} = 1 - m\ell - (g + j\ell)(h + ik),$$
$$\sigma_{12} = -(h + ik)(s(a + jb) + ejn) - s(mb + f) - men,$$
$$\sigma_{21} = -d - (g + j\ell)(p + it) - q\ell,$$
$$\sigma_{22} = 1 - (p + it)(s(a + jb) + ejn) - s(c + qb) - e(r + qn).$$

The submatrix S_{22} of the inverse S is given by the inverse of matrix Σ; and, therefore, the elements s_{66}, s_{67}, s_{76}, and s_{77} of S are given by:

$$s_{66} = \sigma_{22}/D$$
$$s_{67} = \sigma_{12}/D$$
$$s_{76} = \sigma_{21}/D$$
$$S_{77} = \sigma_{11}/D,$$

where $D = \sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}$.
The submatrix S_{12} can be calculated from,

$$S_{12} = -(A_{11}^{-1}A_{12})\Sigma^{-1},$$

which yields the following elements of S:

$$s_{16} = s_{76}s$$
$$s_{17} = s_{77}s$$
$$s_{26} = s_{76}e$$
$$s_{27} = s_{77}e$$
$$s_{36} = s_{66}i(g + j\ell) + s_{76}i(s(a + jb) + ejn)$$
$$s_{37} = s_{67}i(g + j\ell) + s_{77}i(s(a + jb) + ejn)$$
$$s_{46} = s_{66}(g + j\ell) + s_{76}(s(a + jb) + ejn)$$
$$s_{47} = s_{67}(g + j\ell) + s_{77}(s(a + jb) + ejn)$$
$$s_{56} = s_{66}\ell + s_{76}(sb + en)$$
$$s_{57} = s_{76}\ell + s_{77}(sb + en).$$

The next step is to calculate the submatrix,

$$S_{21} = -\Sigma^{-1}(A_{21}A_{11}^{-1}),$$

which gives the following elements of S:

$$s_{61} = s_{66}(f + (h + ik)(a + jb) + mb) + s_{67}(c + (p + it)(a + jb) + qb)$$
$$s_{71} = s_{76}(f + (h + ik)(a + jb) + mb) + s_{77}(c + (p + it)(a + jb) + qb)$$
$$s_{62} = s_{66}(n(j(h + ik) + m)) + s_{67}(r + n(j(p + it) + q))$$
$$s_{72} = s_{76}(n(j(h + ik) + m) + s_{77}(r + n(j(p + it) + q))$$
$$s_{63} = s_{66}k + s_{67}t$$
$$s_{73} = s_{76}k + s_{77}t$$
$$s_{64} = s_{66}(h + ik) + s_{67}(p + it)$$
$$s_{74} = s_{76}(h + ik) + s_{77}(p + it)$$

$$s_{65} = s_{66}(j(h + ik) + m) + s_{67}(j(p + it) + q)$$
$$s_{75} = s_{76}(j(h + ik) + m) + s_{77}(j(p + it) + q).$$

Before calculating S_{11} it is convenient to calculate the matrix,

$$V = (A_{11}^{-1}A_{12})\Sigma^{-1}(A_{21}A_{11}^{-1}) = -(A_{11}^{-1}A_{12})S_{21}.$$

The elements of this matrix are given by:

$$v_{1\alpha} = s_{7\alpha}s$$
$$v_{2\alpha} = s_{7\alpha}e$$
$$v_{3\alpha} = i(g + j\ell)s_{6\alpha} + i(s(a + jb) + ejn)s_{7\alpha}$$
$$v_{4\alpha} = (g + j\ell)s_{6\alpha} + (s(a + jb) + ejn)s_{7\alpha}$$
$$v_{5\alpha} = \ell s_{6\alpha} + (sb + en)s_{7\alpha}$$
for $\alpha = 1, \ldots, 5$.

The matrix S_{11} is then given by,

$$S_{11} = A_{11}^{-1} + V,$$

giving,

$$S_{11} = \begin{bmatrix} 1 + v_{11} & v_{12} & v_{13} & v_{14} & v_{15} \\ v_{21} & 1 + v_{22} & v_{23} & v_{24} & v_{25} \\ i(a + jb) + v_{31} & ijn + v_{32} & 1 + v_{33} & i + v_{34} & ij + v_{35} \\ a + jb + v_{41} & jn + v_{42} & v_{43} & 1 + v_{44} & j + v_{45} \\ b + v_{51} & n + v_{52} & v_{53} & v_{54} & 1 + v_{55} \end{bmatrix}.$$

All the elements of the array S have now been evaluated.

Network Thermodynamics in Biology and Ecology: An Introduction

DONALD C. MIKULECKY

Department of Physiology and Biophysics, Medical College of Virginia, Virginia Commonwealth University, Box 551, MCV Station, Richmond, Virginia 23298-0001, USA

Introduction

The evolution of thermodynamic thinking is a curious one (Mikulecky 1983a). The original focus of attention on equilibrium states seemed to approach the dynamic physical world backwards. In the light of recent advances in dynamic systems theory, one could argue that the proper perspective is to view equilibria as the end points of dynamic events. However, tradition and history are difficult to overcome, and a scientist usually is trained in classical (equilibrium) thermodyamics with a possible smattering of nonequilibrium thermodynamics. One of the most outspoken critics of this approach is Clifford Truesdell (1969). In addition I. Prigogine and his associates have extended thermodynamic reasoning into realms far beyond the scope of most introductory courses, with a view towards describing the complexities of living systems and their dynamic states (Allen, this volume).

It therefore may seem a somewhat radical proposal to suggest that network thermodynamics be considered as the next logical step in the evolution of thermodynamic thinking. But, this proposal may be even truer than the beginning student of this new discipline first recognizes. In fact, after becoming more familiar with the content of network thermodynamics, he will most likely come to regard the approach as revolutionary rather than evolutionary. In this introduction I outline a broad overview of the topic, with the aim of unifying some old, familiar concepts with new ones and of putting this amalgam into a new, and more rational framework. The examples used have some of the "flavor" of problems studied in oceanography but they are obviously oversimplified. They are meant more to illustrate the methods of network thermodynamics rather than to portray ecological systems. The aim of this writer is to stimulate future collaborative applications focussing on biological problems. The suggestion that network thermodynamics is the next step in the development of biological thermodynamics is not meant to be taken lightly.

The paths of nonequilibrium thermodynamics, a significant component of network thermodynamics, and ecology crossed some time ago. Odum's now classical paper on energy conversion (Odum and Pinkerton 1955) and Kedem and Caplan's (1965) treatment of the degree of coupling and maximum efficiency have many ideas in common. The Kedem and Caplan approach has been often applied to problems in biophysics (Caplan and Essig 1983). Recent progress by Peusner (1984b), who uses differential geometry to extend network thermodynamics into the nonlinear domain, has revealed that Kedem and Caplan's degree of coupling is a geometric invariant in nonlinear dissipative systems. Both Odum's and Caplan's approaches can now be seen as special cases in the unifying framework of network thermodynamics (Peusner 1983).

It is not unusual for similar ideas to arise independently in different areas of biology, or in different fields of science in general. Periodically, unifying concepts will originate to bring these ideas together, and these unifying ideas make up the realm of interdisciplinary research. Modern dynamic systems theory is a good example of such a unifying, interdisciplinary thrust that incorporates areas of mathematics such as topology and differential geometry and serves as a powerful new approach to problems in a number

of traditional disciplines. Network thermodynamics is closely wedded to dynamic systems theory by virtue of the fact that the principles of traditional thermodynamics were first created to describe the behavior of so-called "simple" systems, i.e., homogeneous systems, or at least systems with a minimum of organization and structure. This limitation should suggest that thermodynamics can be broadened and its applications extended to complex, hierarchical systems by invoking the mathematics of organization, *analysis situs* (topology) (Veblen 1931).

Topological methods such as graph theory have already been employed in population dynamics (Lewis 1977) and in the domain of holistic biology (MacDonald 1983). One example of these approaches was introduced at this conference (presentation by P. A. Lane). The network thermodynamic approach, although closely related to graph theoretical methods, is a more comprehensive discipline which unites them with the existing body of thermodynamic and kinetic knowledge. The following brief introduction will but scratch the surface of network thermodynamics, but it is intended to encourage further examination of the potential role of this theory in ecology.

The Structure of Network Thermodynamics

Network thermodynamics is based on certain generalizations of the manner in which circuit theorists analyze electrical networks. The reader is cautioned that what follows can, at first glance, look like electrical analogy, equivalent circuits and several other treatments already introduced to one or more areas of biology. However, the fact is that these generalizations, though they were inferred from the analysis of electrical networks, need no further reference to electrical circuits to remain valid. Nevertheless, it is a pedagogical convenience to rely on electrical concepts to a much greater extent than is really necessary.

The most obvious remnants of electrical network analysis lie in certain key steps in the approach. First of all, complex systems are viewed as a collection of discrete elements which are interconnected in a certain way. This description applies to any highly organized system, including ecological communities. The many diagrams of the models of oceanographic systems reflect this assumption (Platt et al., 1981). In order to represent a complex system as a network of discrete elements it is helpful to codify the identities of the elements and the interactions among them in schematic fashion. In network thermodynamics, two distinctly separate systems of representation have evolved. One method consists in the use of diagrams called "bond-graphs" (Oster et al. 1971, 1973; Oster and Perelson 1973, 1974; Perelson and Oster 1975; Karnopp and Rosenberg, 1975; Thoma 1975; Breedveld 1984), and the other, which utilizes linear graphs like those in electronics, is the method of Peusner (1984a). Any choice between these representations is somewhat arbitrary, inasmuch as the bond graphs can be translated into linear graphs and back again at will. Linear graphs will be employed in this paper.

There is some question as to how easily bond graphs are understood by most biologists (Mikulecky and Thomas, 1979), but they certainly are finding wide application in other disciplines (Thoma 1975; Breedveld 1984; Gebben 1979). The style of diagrams used in compartmental analysis, the study of chemical reaction kinetics, as well as the symbols introduced by Odum and others into ecology are alternative representations which can, in most cases, be translated into network thermodynamic schematics. However, such diagrammatic mappings are secondary to the ability they lend one to analyze the system in a systematic manner, which is the main contribution of network thermodynamics.

The next step in the network thermodynamic approach is the mathematical codification of the geometry and topology of the system. Describing the geometry involves identifying the classes of elements which are inherent in the generalized expression for power loss or

gain in a system. The power is a product of an effort variable, e, and a flow variable, f, (Oster et al. 1973):

$$P = ef$$

The effort may be expressed as a difference in *potential-like quantities* existing across the element, and the flow as the movement of a certain amount of a fundamental quantity (mass, volume, charge) through the element per unit time. These effort and flow variables are chosen in such a way to obey Kirchhoff's generalized voltage and current laws, KVL and KCL. The current or flow law rests on the conservation of the flowing quantity, i.e., charge, mass or volume in an incompressible fluid system. The voltage or force law is the result of the closure property, namely that the differences in any potential-like quantity around a closed loop sum to zero. This property is of great importance to the structure of network thermodynamics and may be lacking in some systems. One method of insuring that network thermodynamics may be applied to analyze a system is to *define*, if possible, the efforts and flows as quantities which perforce obey KVL and KCL. The absence of these properties from a system will necessitate the use of a more general network approach, such as signal-flow graphs.

Another caution is in order at this point. In most physical dynamical systems the effort is seen to "drive" the flow through an element. This causal link is both convenient and useful. In the application of network thermodynamics to the physiology or biochemistry of single organisms (Mikulecky 1983d) this causal relationship between force or effort and flow is a physical reality. In dealing with populations of organisms, however, the identification of an effort conjugate to a flow will be necessarily a more *formal* exercise. Nevertheless, if done correctly, the formalism will retain all the analytical power it brings to the study of more traditional physical systems. Thus, even if the forces are defined so as to be of no intrinsic physical interest, they nonetheless provide a way of relating local and global (holistic) behaviors in any complex system.

Once the definition of effort and flow has been made, two other state variables are easily defined (Oster et al. 1973), namely the generalized displacement, q,

$$q(t) = q(o) + \int_o^t f(t)dt,$$

and the generalized momentum, p,

$$p(t) = p(o) + \int_o^t e(t)dt.$$

Thus, the amount of material accumulated in a volume element by the flow through its surface becomes a displacement, n,

$$n(t) = n(o) + \int_o^t J(t)dt,$$

where the flow, $J(t)$, is defined as

$$J(t) = \frac{dn(t)}{dt}.$$

One of the simplest formal methods for assigning an effort conjugate to a flow is to insert the flow quantity into a consistent set of constitutive relations for resistance and capacitance. For example, in Fig. 1 the nodes represent two spatially, taxonomically or trophically distinguishable "pools" of biomass. The object between them is formally a "dissipator", which in general has the constitutive relation,

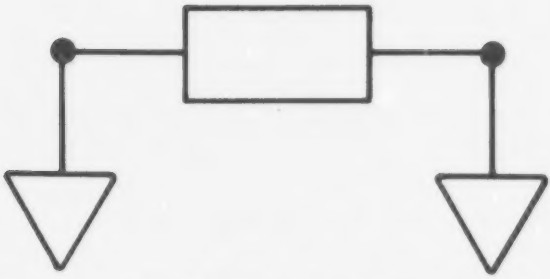


FIG. 1. Two distinguishable "pools" of biomass, separated by a dissipator (rectangle). The capacity of the pools for biomass is represented by the triangular volume capacitors, which relate the amount or mass to the concentration. In this way, the flow through the dissipator is the rate of biomass flow (mass units per unit time), and the conjugate effort is the difference in concentration of biomass between the two pools (mass per unit volume).

$$e = R(f),$$

where R is a nonlinear function. In a more mechanistic system this dissipator could obviously be interpretated as a "frictional" resistance to flow. In the biological realm, however, the interpretation is less clear. If R has an inverse, then an alternative definition becomes

$$f = R^{-1}(e).$$

How do we choose e? The triangular elements connected to each node are the analogs of electrical capacitors to ground and have the dimensions of volume. These capacitors, if linear (constant in volume or capacity), have the constitutive relations,

$$V\frac{de}{dt} = f.$$

Now, if the flow, f, is defined as the amount (of biomass, for example) flowing per unit time,

$$f = \frac{dn}{dt},$$

and the effort as the difference in the biomass/volume between the two nodes, a self-consistent picture emerges. As in any system, physical or otherwise, the nature of the nonlinear constitutive functions, or the values of the linear constants, must be determined empirically.

Constitutive Relations

Once the basic state variables have been defined, there are four pertinent bilaterial constitutive relations among them which define additional categories of network elements (Oster et al. 1973). They are:

capacitance, $C(e, q) = 0$;
inductance, $L(p, f) = 0$;
resistance, $R(e, f) = 0$;
and memristance, $M(p, q) = 0$.

In most biological applications capacitance and resistance are the more commonly encountered elements. However, when inertial effects must be considered, such as in fluid flow, the inductance element is utilized (Rideout and Katra 1969). Any of these constitutive relations may be either linear or nonlinear.

An additional set of elements, the sources, enter into the picture, but closer examination will show that they are special cases of the four basic categories. For example, a linear, passive resistor can be quantified by the slope of a line through the origin on an effort vs flow (or flow vs effort) plot. When this slope becomes zero (infinite) we have the special case of a constant source. A constant flow source maintains a given flow regardless of the value of the effort variable across it. Likewise, a constant effort source always has the same effort across it irrespective of the flow through it.

A more complicated element is the unistor. It is an intrinsically non-reciprocal element in the sense of Onsager and strictly speaking is not defined in terms of the effort and flow variables. An example would be the unidirectional kinetic step in compartmental analysis or in chemical kinetics, and there are many similar processes in ecological models. The constitutive relation for the unistor is:

$$\text{flow} = k \times \text{potential},$$

where k is the rate constant and the potential is the same as that used to define the effort across the element, but measured only on the feed side. If these kinetic steps are always taken in pairs so that netflow = forward flow − backward flow, the result is the same as a dissipator. It is thus easily shown that with the exercise of some care in redefining the efforts in the system, unistors can be eliminated, and any network containing unistors can be made equivalent to an ordinary resistive network with sources (Peusner 1982; Mikulecky 1983b). Thus, the formalisms of compartmental analysis (Mikulecky et al. 1979; Mikulecky 1984; Thakker et al. 1982), chemical kinetics (Wyatt et al. 1980; White and Mikulecky 1982; Mikulecky 1983c; Peusner et al. 1984) and of any others isomorphic with these two categories are readily translated into the form of network thermodynamics. Hence, the following methods of analysis should be applicable to all classes of problems involving dynamic systems.

Network Topology

Even before specifying the constitutive relations for the elements of a network, one can learn much from a knowledge of how the elements are connected to each other — the network topology. The simplest way of encoding topology into diagramatic form is to transcribe the network into a directed linear graph or diagraph. An example of a very simple ecological subsystem which might be suitable for network analysis is shown in Fig. 2a. The corresponding network is shown in Fig. 2b and its associated diagraph in Fig. 2c. This extremely simple example will serve to illustrate the analysis which is generally applicable to all such networks. The topology of the graph is encoded in a set of incidence matrices. The simplest matrix in the set is the node-branch incidence matrix, whose elements are:

$$a_{ij} = \begin{cases} +1 & \text{if branch } i \text{ is incident on node } j \text{ and directed away from it} \\ 0 & \text{if branch } i \text{ is not incident on node } j \\ -1 & \text{if branch } i \text{ is incident on node } j \text{ and directed into it.} \end{cases}$$

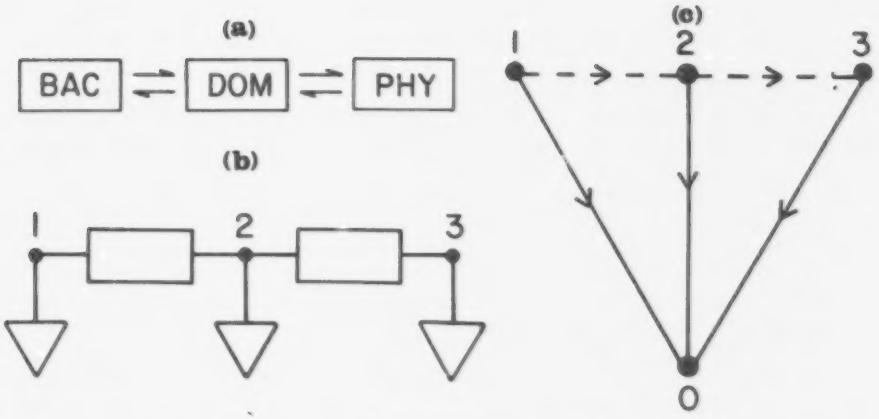


FIG. 2. (a) A simplified ecological subsystem depicting the partitioning of carbon between dissolved organic matter (DOM), bacteria (BAC), and phytoplankton (PHY). (b) The network for the system in (a) with node 1 representing the phytoplankton carbon pool, node 2 the dissolved organic carbon pool and node 3 the bacterial carbon pool. The capacity of each pool is symbolized by the triangular shaped elements which corresponds to constitutive relations relating the amount of carbon to the "volume" of the pool. As carbon flows from one pool to another the carbon "concentration" (mass of carbon per unit volume) changes with inflow or efflux. The effort variable conjugate to the flow of carbon is, therefore, the carbon concentration in each pool. The choice for the "volume" is flexible. For example, it could be the biomass of the pool, as long as a biomass per unit volume were known. (c) The diagraph for the network in (b). The capacitors all connect to a common reference node labeled "zero."

Thus, the node-branch incidence matrix for the graph in Fig. 1c is,

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

where the five branches (in the order 12, 23, 10, 20, and 30 from left to right) identify the columns, and the four nodes (in the order 1, 2, 3, and 0 from top to bottom) define the rows. Not only is this a perfectly fine analytical representation of the topology of the graph, but it also has other interesting properties. For example, multiplying the vector of branch flows from the left by the incidence matrix recapitulates the KCL:

$$\underline{A} \cdot \bar{f} = \bar{0},$$

or

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} J_{12} \\ J_{23} \\ J_{10} \\ J_{20} \\ J_{30} \end{bmatrix} = \begin{bmatrix} J_{12} + J_{10} \\ -J_{12} + J_{23} + J_{20} \\ -J_{23} + J_{30} \\ -J_{10} - J_{20} - J_{30} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Likewise, when the transpose of the node branch matrix multiplies the node potentials the product vector is the branch efforts—a version of KVL:

$$\underline{A}^T \cdot \bar{v} = \bar{e},$$

or

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_0 \end{bmatrix} = \begin{bmatrix} v_1 - v_2 \\ v_2 - v_3 \\ v_1 - v_0 \\ v_2 - v_0 \\ v_3 - v_0 \end{bmatrix} = \begin{bmatrix} e_{12} \\ e_{23} \\ e_{10} \\ e_{20} \\ e_{30} \end{bmatrix}$$

There are two other incidence matrices which alternatively can be used to produce KCL and KVL, but we need not develop them here. They are the loop-branch and the cut-set incidence matrices and are discussed in detail elsewhere (Mikulecky 1984; Athans et al. 1974; Desoer and Kuh 1969).

Tellegen's theorem

These alternative manifestations of the fact that the effort and flow were defined so as to obey KCL and KVL lead to a very important result. If we use the KVL to express the scalar product between the effort and flow vectors as,

$$\bar{e} \cdot \bar{f} = (\underline{A}^T \bar{v}) \cdot \bar{f},$$

then by a simple vector theorem this scalar product can be shown to be zero:

$$\bar{e} \cdot \bar{f} = (\underline{A}^T \bar{v}) \cdot \bar{f} = \bar{v}^T \cdot (\underline{A}\, \bar{f}) = \bar{v}^T \cdot \bar{0} = 0.$$

This statement is a form of Tellegen's theorem, a power conservation law guaranteeing that power in minus power consumed is zero. In addition one observes that, in general, effort and flow vectors obey Kirchhoff's laws are also orthogonal. Not only this, but the effort and flow vectors from different networks are orthogonal as long as they have the same topology! This result as it applies to a given network is Tellegen's power conservation theorem, and its extension to different networks with identical topology is known as the quasi-power theorem (Oster et al. 1971, 1973, 1974; Perelson et al. 1975; Peusner 1984a; Mikulecky et al. 1979). One rather striking corrollary of this theorem is Onsager's requirement for reciprocal coupling in linear, multiport networks (Peusner 1984b). Notice that these results are irrespective of the nature or identity of the network elements in the branches, i.e., they require no knowledge of the constitutive relations (geometry). They are purely topological properties of networks that obey Kirchhoff's laws. The central role of Kirchhoff's laws in network thermodynamics now should be clear.

The Relationship of Network Thermodynamics to Nonequilibrium Thermodynamics: the Importance of Tellegen's Theorem

The resistors and sources are the only constituitive elements needed to describe any system in steady state. The capacitors and inductors play a role only in non steady state (dynamical) situations. If the passive dissipative elements are considered as a single unit and connected to sources via pairs of external network terminals called "ports" (Fig. 3), the system can be simplified to a network of a single resistor connected to a source (Fig. 4). Now it becomes easy to visualize the relationship between the dissipation function of nonequilibrium thermodynamics and Tellegen's theorem. Notice that if, by

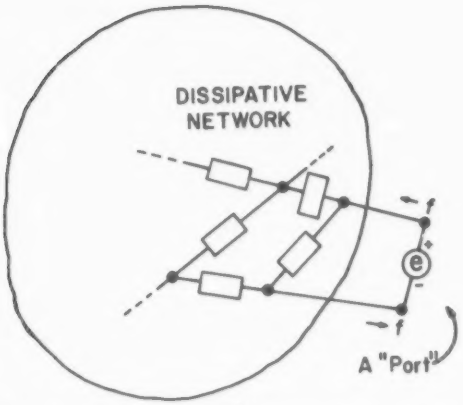


FIG. 3. The demarcation between dissipative elements and energizing sources distinguishes between the orthogonal flows and efforts of Tellegen's theorem and the dissipation function of nonequilibrium thermodynamics.

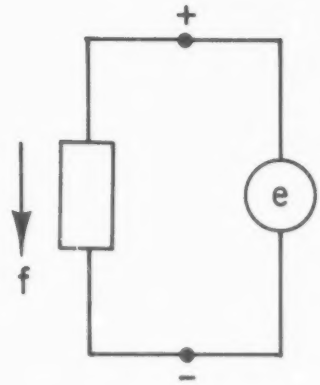


FIG. 4. A single resistor and its energizing source is the simplest possible example of how sign conventions lead to flow-effort orthogonality through power conservation.

convention, the positive direction of flow is defined to proceed from the positive node through the resistor to the negative node, the same quantity of flow must necessarily pass in the negative direction back through the source. Thus Tellegen's theorem holds. In more complicated networks, this same sign convention prevails, and Tellegen's theorem takes on the guise of a power conservation law. However, the inner passive network is the part usually described by nonequilibrium thermodynamics. By adding the power supplied by the sources to the description of dissipation among the elements of the system, the orthogonality of effort and flow vectors becomes evident, and the metric structure is shown to be a property of the power space (Peusner 1984b). These results are far from trivial an are the basis for the earlier statements advocating the role of network thermodynamics in extending thermodynamic thinking. Prior to Peusner's demonstration the thermodynamics of dynamic systems lacked a "yard-stick" (metric) to measure the distance between states.

Multiport Dissipators for Coupled Processes

A large school of fish moving from one point in the ocean to some distant location obviously carry with them carbon, nitrogen and biomass. Water movements also transport small organisms, sediments and particulate matter. These flows are "coupled", and to ignore the coupling is to ignore valuable descriptive and analytical information about the system.

A very primitive example of a multiport dissipator representing such coupled processes is shown in Fig. 5a and b. The flows of nitrogen, carbon and biomass make up a three-port element with the linear constitutive relations defined as,

$$\bar{J} = \begin{bmatrix} J_b \\ J_c \\ J_n \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} e_b \\ e_c \\ e_n \end{bmatrix}$$

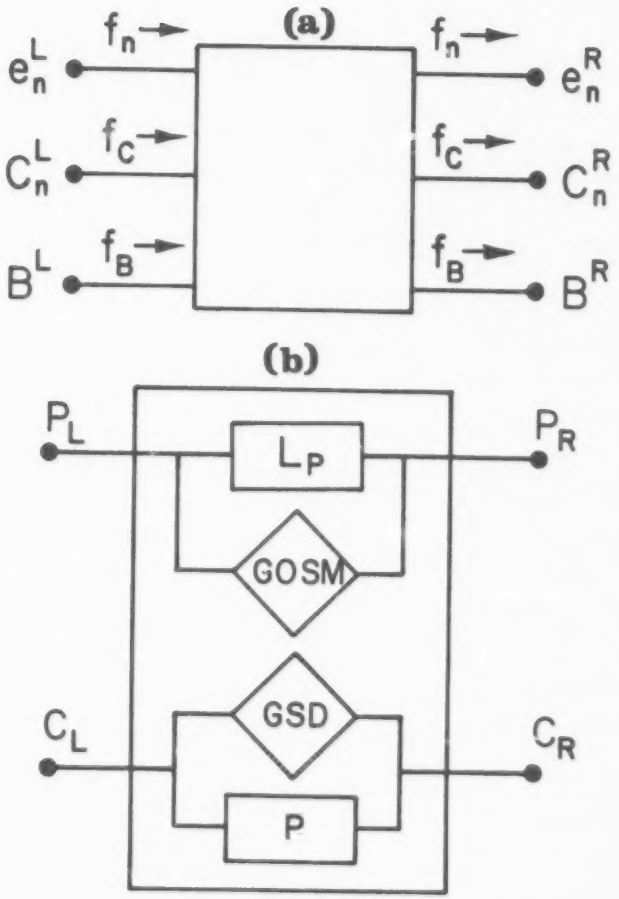


FIG. 5. Multiport resistors for coupled flows: (a) A three-port resistor for coupled carbon, nitrogen and biomass flow. Each pool has its own capacitor connected to the multiport. These elements are easily interconnected. (b) A two-port resistor for coupled solute and convective flow. The rectangular elements inside the "black box" depict dissipators linking the conjugate efforts and flows, while the diamond-shaped elements in parallel depict the coupled flow of volume to concentration difference of solute (GOSM), and the coupled flow of solute due to the pressure differences which produce convective volume flow (GSD) (Thomas and Mikulecky 1978).

The constitutive relations are assumed to be linear only for purposes of illustration. The general case could have been expressed as,

$$\bar{J} = \begin{bmatrix} J_b \\ J_c \\ J_n \end{bmatrix} = \begin{bmatrix} L_b & (e_b, e_c, e_n) \\ L_c & (e_b, e_c, e_n) \\ L_n & (e_b, e_c, e_n) \end{bmatrix}$$

The example is almost trivial, since the flow of carbon and nitrogen in a school of swimming fish can be expressed as,

$$J_n = k_n J_b$$
$$J_c = k_c J_b,$$

and the system is seen to be completely coupled. In other words, it reduces to

$$\bar{J} = \begin{bmatrix} J_b \\ J_c \\ J_n \end{bmatrix} = \begin{bmatrix} 1 \\ k_n \\ k_c \end{bmatrix} L e_b .$$

However, other situations can easily be envisioned, such as in the study of Joiris et al. (1982) wherein the relationship of flow of carbon to that of biomass or energy is much more subtle, and the representation as a coupled multiport system becomes not only more interesting, but actually necessary to achieve a complete system description.

The convective flow of suspended particles is not unlike the convection of solutes in physiological systems. As there are many examples of how multiports facilitate the analysis of such systems (Oster et al. 1973; Mikulecky et al. 1979; Thomas and Mikulecky 1978), the subject will not be elaborated here.

Solutions to dynamical network

The most useful version of network analysis involves the use of cut sets and their corresponding incidence matrices to solve dynamical networks. These methods are developed in detail elsewhere (Mikulecky 1984; Athans et al. 1974; Desoer and Kuh 1969) and will not be repeated here. To illustrate the manner in which topology and geometry combine to produce a solution to system dynamics, consider the network in Fig. 1b to be at steady state. It will contain only two "dissipative" elements as represented by the rectangular boxes. A very general solution to this or any similar problem can always be generated directly by combining the general forms of KCL and KVL with the constitutive relations. Allowing for the presence of sources, a generalized dissipative branch can be represented as in Fig. 6. This branch has the constitutive relation,

$$J_{ik} = L_{ik} e_{ik} - L_{ik} S_{ik} + I_{ik} ,$$

where J_{ik} is the total flow through the branch
L_{ik} is the conductivity of the dissipator
e_{ik} is the effort across the branch
S_{ik} is the effort of the source (acting in series)
I_{ik} is the (parallel) flow through the source.

By defining a diagonal conductance matrix, $\underline{L}$ the constitutive laws can be expressed in matrix-vector form,

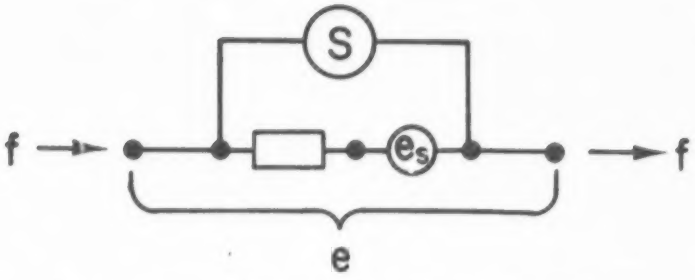


FIG. 6. The general dissipative branch with a dissipator and force source in series and a flow source in parallel.

$$\underline{L} = \begin{bmatrix} L_{11} & 0 & 0 & \text{----} & 0 \\ 0 & L_{22} & 0 & \text{----} & 0 \\ 0 & 0 & L_{33} & \text{----} & 0 \\ 0 & 0 & 0 & \text{----} & Lnn \end{bmatrix},$$

where n is the number of branches.

The constitutive relations then appear as,

$$\bar{J} = \underline{L}\bar{e} - \underline{L}\bar{S} + \bar{I}.$$

Multiplying from the left by the incidence matrix and using KCL gives:

$$\bar{A}\,J = \underline{AL}\bar{e} - \underline{AL}\bar{S} + \underline{A}\bar{I} = \bar{0}.$$

Rearranging, using KVL, and substituting for the effort vector leads to,

$$\underline{ALA}^T\bar{v} = \underline{AL}\bar{S} - \underline{A}I.$$

Define the node-branch admittance matrix, $\underline{Y}$, as:

$$\underline{Y} = \underline{ALA}^T.$$

Because $\underline{Y}$ can be shown to possess an inverse, the node potentials may be obtained in terms of the source strengths,

$$\bar{v} = \underline{Y}^{-1}\underline{AL}\bar{S} - \underline{Y}^{-1}\underline{A}\bar{I}.$$

Now by reversing the procedure, the efforts and flows are obtained from the node potentials, and the steady-state problem has been solved. Notice that at no point did a causal relationship between effort and flow need to be invoked. However, this example rests on the *linearity* of the constitutive relations between effort and flow. This constraint can be relaxed and the method modified to handle nonlinearities, but not without difficulty.

It is a simple matter to introduce more graph theory and to use cut set analysis (this method likewise depends on topological incidence, KCL, KVL and the constitutive relations) to solve linear, nonsteady state networks with capacitance and inductance (Athans et al. 1974; Desoer and Kuh 1969; Mikulecky 1984; Chua 1969).

Computer simulation of networks

The use of computer simulation in network thermodynamics has become an increasingly

useful tool in the experimental interpretation and design of systems (Mikulecky et al. 1979; Thakker et al. 1982; Wyatt et al. 1980; White and Mikulecky 1982; Mikulecky 1983b; 1983c). Since numerical simulation has been employed in so many demonstrations of the applications of network thermodynamics to physiological, biochemical and pharmacokinetic systems, it will not be presented again here. Rather, its utility, economy and ease of mastery will be demonstrated in some future simulations of ecological systems. Hopefully, this exercise will involve the colaboration of the experimental ecologists themselves (Mikulecky 1983d). The continual improvement of the already powerful software for simulating electrical networks is anticipated for some time to come. Likewise, the eventual creation of similar programs with a more biological syntax and with "canned" biological elements is inevitable.

The ease with which this paradigm could become a common language for experimental and theoretical biology should not be ignored (Mikulecky 1983d; Jaynes 1976). The collective consciousness of a community of scientists may indeed be the mode in which future scientific progress will be made. Problems in ecology are complex, yet it is critical that they be understood and solved. For this reason, ecology is a good area in which to attempt to achieve a collective consciousness.

The systematic nature of problem solving inherent in network thermodynamics should serve as a building block for a more general theory that will incorporate the stochastic nature of ecological systems (Tan and Wyatt 1984), as well as the variability of their topology. There are certainly enough aspects of ecological systems which fall under the purview of the conservation laws (KCL and KVL) and which can be fruitfully examined using these powerful methods.

References

Athans, M., M. L. Dertouzos, R. N. Spann, and S. J. Mason. 1974. Systems, networks, and computation: multivariable methods. McGraw-Hill, New York, NY.

Breedveld, P. C. 1984. Physical systems theory in terms of bond graphs. Ph.D. thesis, Enschede.

Caplan, S. R., and A. Essig. 1983. Bioenergetics and linear nonequilibrium thermodynamics: the steady state. Harvard, Cambridge, MA.

Chua, L. O. 1969. Introduction to nonlinear network theory. McGraw-Hill, New York, NY.

Desoer, C. A., and E. S. Kuh. 1969. Basic circuit theory. McGraw-Hill, New York, NY.

Gebben, V. D. 1979. Bond graph bibliography. J. Franklin Inst. 308: 361–369.

Jaynes, J. 1976. The origin of consciousness in the breakdown of the bicameral mind. Houghton Mifflin, Boston, MA.

Joiris, C., G. Billen, C. Lancelot, M. H. Daro, J. P. Mommaerts, A. Bertels, M. Bossicart, J. Nijs, and J. H. Hecq. 1982. A budget of carbon cycling in the Belgian Costal Zone: relative roles of zooplankton, bacterioplankton and benthos in the utilization of primary production. Neth. J. Sea Res. 16: 260–275.

Karnopp, D., and R. Rosenberg. 1975. System dynamics: a unified approach. Wiley, New York, NY.

Kedem, O., and S. R. Caplan. 1965. Degree of coupling and its relation to efficiency of energy conversion. Trans Faraday Soc. 61: 1897–1911.

Lewis, E. R. 1977. Network models in population biology. Springer, New York, NY.

MacDonald, N. 1983. Trees and networks in biological models. Wiley, New York, NY.

Mikulecky, D. C. 1983a. Unpublished lecture notes for a course on network thermodynamics.

1983b. Onsager's reciprocity in reaction-diffusion systems: a network thermodynamic approach. Va. Acad. Sci 34: J. 188.

1983c. A network thermodynamic approach to the Hill–King and Altman approach to kinetics: computer simulation, p. 257–282. *In* M. A. Dinno, A. G. Calahan, and T. C. Rozzell [ed.] Membrane biophsycs II: physical methods in the study of cellular biophysics.

1983d. Network thermodynamics: a candidate for a common language for theoretical and experimental biology. Am. J. Physiol. 245: R1–R9.

1984. Network thermodynamics: a simulation and modeling method based on the extension of thermodynamic thinking into the realm of highly organized systems. Math. Biosci. 72: 157–179.

MIKULECKY, D. C., AND S. R. THOMAS. 1979. Some network thermodynamic models of coupled, dynamic physiological systems. J. Franklin Inst. 308: 309–326.

MIKULECKY, D. C., E. G. HUF, AND S. R. THOMAS. 1979. A network thermodynamic approach to compartmental analysis: Na^+ transients in frog skin. Biophys. J. 25: 87–106.

MIKULECKY, D. C., W. A. WIEGAND, AND J. S. SHINER. 1979. A simple network thermodynamic method for series-parallel coupled flows. I. The linear case. J. Theor. Biol. 69: 471–510.

ODUM, H. T., AND R. C. PINKERTON. 1955. Time's speed regulator: the optimum efficiency for maxim power output in physical and biological systems. Am. Sci. 43: 331–343.

OSTER, G. F., AND A. S. PERELSON. 1974. Chemical reaction dynamics. Part I. Geometrical structure. Arch. Ration. Mech. Anal. 55: 230–274.

1973. Systems, Circuits and Thermodynamics. Isr. J.Chem. 11: 445–478.

OSTER, G., A. PERELSON, AND A. KATCHALSKY. 1971. Network thermodynamics. Nature 234: 393–399.

OSTER, G. F., A. S. PERELSON, AND A. KATCHALSKY. 1973. Network thermodynamics: dynamic modeling of biophysical systems. Q. Rev. Biophys. 6: 1–134.

PENFIELD, P., JR., R. SPENCE, AND S. DUINKER. 1970. Tellegren's theorem and electrical networks. MIT Press, Cambridge, MA.

PERELSON, A. S., AND G. F. OSTER. 1975. Chemical reaction dynamics. Part II. Reaction networks. Arch. Ration. Mech. Anal. 57: 31–98.

PEUSNER, L. 1982. Global reaction-diffusion coupling and reciprocity in linear asymmetric kinetic networks. J. Chem. Phys. 77: 5500–5507.

1983. Hierarchies of irreversible energy conversion systems: a network thermodynamic approach. J. Theor. Biol. 102: 7–39.

1984a. Studies in network thermodynamics. Entropy, Lincoln, MA.

1984b. Why are Onsager's equations reciprocal? The topology and Euclidean geometry of fluctuation dissipation space. (submitted for publication).

PEUSNER, L., D. C. MIKULECKY, S. R. CAPLAN, AND B. BUNOW. 1984. A unified graphical approach to dynamic systems: network thermodynamics applied to Hill and King–Altman diagrams in reaction-diffusion kinetics. (In press)

PLATT, T., K. H. MANN, AND R. E. ULANOWICZ [ed.] 1981. Mathematical models in biological oceanography. Monographs on Oceanographic Methodology No. 7, The UNESCO Press, France. 156 p.

RIDEOUT, V. C., AND J. A. KATRA. 1969. Computer simulation study of the pulmonary circulation. Simulation 12: 239–245.

ROSEN, R. 1970. Dynamical system theory in biology. Wiley, New York, NY.

TAN, H., AND J. L. WYATT. 1984. Nonlinear network dynamics in the presence of thermal noise. IEEE Trans. Circuits and Stystems. (In press)

THAKKER, K. M., J. H. WOOD, AND D. C. MIKULECKY. 1982. Dynamic simulation of pharmacokinetic systems using the electrical circuit analysis program SPICE2. Comp. Prog. Biomed. 15: 61–72.

THOMA, J. U. 1975. Introduction to Bond Graphs and Their Applications. Pergamon, New York, NY.

THOMAS, S. R., AND D. C. MIKULECKY. 1978. A network thermodynamic model of salt and water flow across the kidney proximal tubule.

TRUESDELL, C. 1969. Rational thermodynamics. McGraw-Hill, New York, NY.

VEBLEN, O. 1931. Analysis Situs. Providence. Am. Math. Soc.

WHITE, J. C., AND D. C. MIKULECKY. 1982. Application of network thermodynamics to the computer modeling of the pharmacology of anticancer agents: a network model for methotrexate action as a comprehensive example. Pharmacol. Ther. 15: 251–291.

WYATT, J. L., JR., D. C. MIKULECKY, AND J. A. DESIMONE. 1980. Network modeling of reaction-diffusion systems and their numerical solution using SPICE2. Chem. Eng. Sci. 35: 2115–2128.

V. INFORMATION THEORY

The Statistical Basis of Ecological Potentiality

MICHAEL CONRAD

Departments of Computer Science and Biological Sciences, Wayne State University, Detroit, MI 48202, USA

Introduction

I have been asked to discuss adaptability theory and how it relates to statistical mechanics, biological oceanography, and marine production. Adaptability theory is a statistical theory in the sense that it deals with ecological and evolutionary problems in terms of ensembles of systems and measures of uncertainty about these ensembles. The theory also has a strong thermodynamic undercoat. Adaptability is the ability to absorb or dissipate disturbances. Issues of thermodynamic efficiency enter. However, the techniques used in adaptability theory are actually much different than those used in either statistical mechanics or thermodynamics. Biological systems are too diverse and too exquisitely constrained to admit of simple averaging procedures. It is the diversity of structure and behavior which is really significant. In fact, the techniques of adaptability theory draw more from information and systems theory than they do from statistical mechanics.

It would be duplicative to review adaptability theory in this paper. A number of review articles have already been presented (Conrad 1972a, 1975, 1976a, 1976b). A thorough review can be found in my recent book (Conrad 1983). What I wish to do here is to suggest the possible significance of adaptability theory for oceanographic problems. Like many of the participants in this conference I am not enormously well versed in the details of oceanography. Nevertheless, I am going to take up Dr. Uianowicz's suggestion and use the problem of marine production as an example, although with the limited intention of making a methodological point rather than a thoroughgoing analysis. The point I wish to make is that the concept of potentiality is essential for a proper description of ecological systems, and that this concept is captured more adequately by the formalism of adaptability theory than by dynamical models patterned after the successes of classical physics.

Thumbnail Review of Adaptability Theory

The fundamental quantities in adaptability theory are measures of behavioral uncertainty, such as the following:

(1) $H(\omega^*)$ = behavioral uncertainty of the environment (ω^* is the environmental transition scheme).

(2) $H(\hat{\omega})$ = potential behavioral uncertainty of the biotic part of the ecosystem ($\hat{\omega}$ is the transition scheme of the biota in the most uncertain tolerable environment).

(3) $H(\hat{\omega}|\hat{\omega}^*)$ = potential behavioral uncertainty of the biota, given the behavior of the environment. (This increases as the ability of the biota to anticipate the environment increases. It also increases as the internally generated uncertainty in the biota increases.)

(4) $H(\hat{\omega}^*|\hat{\omega})$ = potential behavioral uncertainty of the environment given the behavior of the biota. (This increases as indifference of the biota to the environment increases, for example, as organisms live in a smaller region of space. All else being equal, niche breadth decreases when indifference to the environment increases.)

Among these four quantities there obtains the relation,

(1) $H(\hat{\omega}) - H(\hat{\omega}|\hat{\omega}^*) + H(\hat{\omega}^*|\hat{\omega}) \rightarrow H(\omega^*)$.

The lefthand side represents the adaptability of the biotic system of interest. The righthand side represents the actual uncertainty of the environment. The arrow represents a plausible evolutionary tendency of adaptability. All forms of adaptability are costly. Adaptabilities which are never used tend to disappear in the course of evolution. In a culture chamber experiment the uncertainty of the environment is measurable (since it is controllable). The assumption is that systems cultured in more uncertain environments will maintain a higher adaptability.

The relationship between adaptability and environmental uncertainty is actually more complicated than this. Adaptability increases as the difference between behavioral uncertainty and conditioned behavioral uncertainty increases (that is, as $H(\hat{\omega}) - H(\hat{\omega}|\hat{\omega}^*)$ increases). The magnitude of these terms could be individually high, yet the adaptability could remain low. In this case there would be a great deal of biological variability, yet not very much of it would appear as adaptability. However, these reserves of variability can be converted into adaptability in the event of a crisis.

To express the all-important fact of hierarchical organization in ecosystems we can write transition schemes for individual compartments such as populations, organisms, and genomes. The symbol $\hat{\omega}_{ij}$ represents the transition scheme of compartment i at level j in terms of its subcompartments at the next lower level. The uncertainty of the biota can be expressed in terms of a sum of the effective entropies of each compartment,

(2) $H(\hat{\omega}) = \sum_{i,j} He(\hat{\omega}_{ij})$.

Each effective entropy is a sum of conditional and unconditional entropies,

(3) $He(\hat{\omega}_{ij}) = fH(\hat{\omega}_{ij}) +$ conditional terms,

where f is a normalizing coefficient. The unconditional part is the behavioral uncertainty of the compartment considered in isolation, and the conditional parts express the correlation between this uncertainty and the modifiabilities of other compartments. The uncertainty, taken in isolation, will be called the modifiability.

A crucial point is that the adaptability is not the sum of the modifiabilities. If a system is more decentralized (that is, if the parts are more independent), then the adaptability is greater for given observable modifiabilities of the parts. If constraints are added to the system which decrease the conditional entropies, the adaptability must decrease or be compensated by other changes. Other possible changes include enhanced anticipation, increased indifference, or development of new subsystems with high behavioral uncertainty (such as the central nervous system or the immune system). If no such compensation occurs, the niche must narrow or the system must absorb disturbances at the level of population fluctuations. This latter response is an acceptable mode of adaptability only for fast growing organisms (such as microbial flora, phytoplankton, and small organisms in the detrital pathway).[1] Another important contribution to biotic adaptability is the uncertainty of pathways of energy flow, that is, the modifiability of the foodweb structure.

Naturally, adaptability cannot always decrease. This would contradict the overriding fact of life on earth. Many factors control the rise and fall of adaptability. When the

[1]The ability to enter into a spore stage and to switch to a rapidly reproducing stage is a property of individual organisms whose selective value is ultimately based on the fact that some individuals with this strategy survive. With this in mind, it seems legitimate to refer to the resulting dynamics as a populational mode of adaptability.

adaptability of an ecosystem falls, it becomes crisis prone. When a crisis subsequently occurs, a series of changes are instigated which results in the renewal of the adaptability structure.

Adaptability and Dynamics

The transition schemes $\hat{\omega}$ and ω^* are defined on states of the biota and the environment, respectively. In general, transition schemes have a deterministic component (such as the life cycle of the organism or the cycle of the year). They also have an indeterminate component (for example, connected with mutation or unpredictable weather patterns). We do not assume that the biotic transition scheme is known. In fact, I do not assume that it exists as an immutable law of development of a biological system, or even that such immutable laws of development exist. The important point is that the biota and environment have a repertoire of possible modes of behavior. $H(\hat{\omega})$ and $H(\omega^*)$ measure the size of these repertoires at any given time.

The connection between the adaptability formalism and the usual variables of biology is through the decomposition into compartments at different levels of organization. Populations may be described in terms of locations and numbers of organisms (patchiness), organisms in terms of physiological states, genes in terms of base sequences, and so forth. While these identifications are conceptually important, they are not all practical for oceanography. The types of properties which can be measured are physical properties of the environment (including variability), attributes of patches, diversity of biotic components, types of taxa present, and perhaps some indices of genetic and physiological variability. How adaptability theory relates to such observables of oceanography will be considered shortly.

How does adaptability theory relate to dynamical models of the type usually used in biology? An example of a dynamical model is provided by the famous Lotka–Volterra equations,

$$(4)\quad \frac{dN_i}{dt} = N_i\Big(\sum_{j=1}^{m} a_{ij}N_j + a_{i\infty}\Big),$$
$$\vdots \qquad\qquad \vdots$$

where N_i is species i, N_j is species j, and i runs from 1 to m, the number of species. It is possible to translate transition schemes into a dynamical language which admits forms of the above type as special cases. But the interpretation of such models is very different in the adaptability theory framework.

The Lotka–Volterra model is completely deterministic. We ignore here the well-known simplifications and unrealities of models of this type. For the sake of granting the best possible position to advocates of this type of model, we can suppose that it provides an accurate description of a set of biological populations over a meaningful period of time. According to adaptability theory, this is possible only if these dynamics are protected from environmental disturbance. But this requires other dynamics, which are either indeterminate or unstable in an acceptable way. The instabilities of the surrounding dynamics would protect the determinate dynamics. Thus the dynamics of biological systems is not viewed as an independent, immutable thing, in the sense that Newton's law is viewed as immutable within the range of a well-defined approximation (low velocities, reasonably large size). The dynamics are products of evolution, essentially phenotypic traits, always subject to further evolutionary development. If some level of biological dynamics appears immutable, it is only because it is protected by a rich adaptability structure. It is the

potentiality inherent in the adaptability structure, especially the evolutionary potentiality, which is ultimately responsible for the appearance of immutability of the subsystem. But the dynamics of a system with true evolutionary potentiality cannot be immutable.[2]

The Gene Engineered Ocean[3]

Marine biota often have a patchy organization (Legendre and Demers 1984). This is the case for phytoplankton, zooplankton, and even for occupants of higher trophic levels (such as schools of fish). Many specific aggregation mechanisms contribute to the morphogenesis of these organisms. Undoubtedly, patchy organization plays a number of roles. For example, the biological oceanographers have shown that vertical mixing of phytoplankton due to various convection currents in the ocean is a critical factor in marine production. Such flows bring the phytoplankton into contact with photons near the water surface, then into contact with nutrients nearer to the ocean bottom or to the pycnocline. Obviously, the morphology of the organisms and their colonies must to some extent have evolved to utilize the various convective and other physical flow processes in the ocean.

A patchy structure is also a highly decentralized form of organization. Specifying the state (or behavior) of any one patch does not necessarily give very much information about other patches. The independence terms, $H(\hat{\omega}_{\text{patch}}|\hat{\omega}_{\text{other patches}})$, are relatively large, perhaps even comparable to the modifiability term, $H(\hat{\omega}_{\text{patch}})$. In this regime of relatively autonomous patches the adaptability can be high, even if the total observable modifiability of the system is low. In alternative language, the cross-correlation among the components is low. If the cross-correlation were high, the total variability would be low, even if each component exhibited a high degree of changeability.

To sort out the relative values of adaptability and specialization I will adopt the rather fanciful posture of a gene engineer who has become emboldened to design a marine ecosystem. Table 1 lists levels of organization which must be considered and the ir associated modes of adaptability.

TABLE 1. Levels of adaptability in a marine ecosystem.

Level	Mode of adaptability
Marine community	Routability of energy and materials
Populations (phytoplankton, zooplankton, fish, . . .)	Culturability (in terms of number of patches)
	Topographic plasticity
Patches (phytoplankton patches, zooplankton clouds, schools of fish, . . .)	Culturability (in terms of size of patches)
Organism or cell	Physiological adaptability
Gene	Genetic diversity

[2] I believe that this point of view is not alien to that expressed by some other speakers at the conference. In particular R. Rosen has pointed out that error in nature does not correspond to error in dynamical models. Yet it is error in nature, in the form of mutation, which leads to evolution. R. Margalef (1973) some time ago pointed out the fundamental asymmetry in dynamical ecological models: they allow for extinction, but not for evolution.

[3] I am indebted to L. Legendre, T. Platt, and other members of the SCOR working group on physical–biological interactions for discussions influencing this section. Professor Margalef also made stimulating remarks on this subject during his talk.

Our gene engineer could plausibly begin the design process by arguing that patches are a good idea for three reasons.

(1) A decentralized organization allows for a maximum of adaptability for given modifiabilities of the units (as argued above).

(2) Some degree of internal specialization within the patches is possible, perhaps even some specialization of labor among the components. The morphological structures can contribute to mixing, hence to exploiting both light and submerged nutrients. Patch-based mobility could also serve to effectively disperse the population over the whole exploitable region. (Insofar as such mobility allows for variability of spatial distribution it is a form of adaptability, namely topographic plasticity).

(3) Patchiness provides a natural mechanism for routability of energy and material flow in the community. Uncertainty in the pathways of energy flow is all-important for dissipating the effects of variations in energy input or disturbances affecting the use of energy (Odum 1953; MacArthur 1955; Conrad 1972b; Ulanowicz 1980). In adaptability theory it is not just the multiplicity of pathways which is important, but also the variability of the transfer coefficients (Conrad 1972b, 1983). This advantage of patchiness will undoubtedly be attractive to our gene engineer, since he will certainly be interested in the reliability of higher level production (for example, of fish), as well as primary productivity.

We can assume that our gene engineer is not operating in isolation. He must either be employed by private industry and motivated by maximum profit or by a government planning agency interested in maximizing productivity and predictability. To accommodate such imperatives our engineer might be tempted to utilize a single, highly-efficient genetic strain (or monoculture) of low physiological adaptability. Genetic diversity means a genetic load that is efficiency-reducing in any given environment. Similarly, physiological adaptability means extra inducible genes or extra cytoplasmic components, which are efficiency reducing in any given environment.

The rationale here might be that the decentralized, patchy form of organization would by itself offer sufficient adaptability to absorb or dissipate environmental disturbance. To some extent this is true. But in the absence of any genetic and physiological adaptability the effectiveness of patchiness is very much diminished. If all the patches were genetically and physiologically identical, specifying the behavior of one would, in general, give a lot of information about the behavior of others, that is, the independence would no longer be high. Patchiness would still serve to effectively dissipate local disturbances. Such disturbances will affect the patch on which they directly act, but will not propagate to other patches at the same or higher trophic levels. But the response of this system to global disturbances would be another matter. If all patches were identical and without capacity for physiological adaptation, a global disturbance could have a devastating direct effect on each. Propagation would then be a moot issue. However, if the patches were endowed with genetic and physiological plasticities, some patches might perform well in the face of an environmental alteration. Patchiness would then prevent the propagation of disturbance from damaged patches to undamaged patches. Hence, culturability of patches is most useful when these underlying adaptabilities are present.

So our gene engineer will find, perhaps through sad experience, that efficiency obtained at the expense of draining essential adaptabilities from the community is a pyrrhic gain, at best. If he cannot convince his supervisors of this, the productivity of the seas would inevitably become seriously compromised, much to the discomfort of our human species. Fortunately, there is no way of engineering adaptability out of the community over the long run. If the community survives a catastrophic disturbance, the changes induced by the catastrophe would provide an opportunity for many inefficient forms to prosper,

resulting in more variability, and hence more of the stuff which can be converted into adaptability.

Patterns of Patchiness

Whether the addition of genetic or physiological adaptability will be more effective depends on the nature and time scale of the disturbance. It is a matter of costs and advantages which must be superimposed on the basic formalism (either through experiment or through separate theoretical arguments). These cost functions are determined by the morphology of the organisms, the internal specialization of the patches, as well as by the particular nature and time scale of the disturbance. For example, the cost of culturability should increase allometrically with the size of the organism (or, more precisely, with the thermodynamic cost of reproduction and growth).[4] Once the cost functions are specified, the problem of ascertaining the spectrum of adaptabilities in a system may be approached as an optimization problem.

It is interesting to consider how altering the uncertainty and harshness of the environment might affect marine organization. Adaptability theory suggests the following two conjectures:

(1) Varied, uncertain environments should favor more decentralized forms of organization. Internal organization is too costly to maintain in such environments. Similarly, a high degree of physiological plasticity might be too costly to maintain. Some physiological plasticity, genetic diversity, and a high degree of patchiness should be favored, as patchiness is a thermodynamically inexpensive option.

(2) Mild, relatively certain environments favor patches with higher genetic and physiological plasticity and allow patches with a higher degree of internal organization.

There is an analogy here to the dichotomy between arctic and tropical ecosystems. Arctic systems have a much more patchy structure and are reported both to undergo more pronounced oscillations and to maintain higher production levels (MacArthur 1955). In these systems local demes absorb perturbation relatively independently of one another. The richer environment of tropical systems allows much more elaborate, interconnected structure, with many specialized species. There is an advantage in terms of efficiency. This pattern of organization is possible because sufficient energy is available to invest in relatively more expensive modes of adaptability.

An experimental fact (from my own freshwater microcosm experiments) is that flask ecosystems cultured in a mild, certain environment often do better when switched to a harsh, uncertain environment than do the microcosms initially cultured in a harsh, uncertain environment when switched to a mild, certain environment. The reason is that the mild, certain environment allows for more rapid accumulation of biomass, hence a greater amount of variability which can be converted to adaptability. Systems cultured in harsh, uncertain environments responded to such switches more favorably if they were made later in their development. The reason is that high variability must await the accumulation of biomass in such disturbed systems. I mention this to draw attention to the fact that the temporal aspect of succession is of critical importance to the development of adaptability. Harshness and uncertainty of the environment affect the temporal development of patterns of adaptability, as well as the final pattern which develops.

Finally, our gene engineer should consider the effect of taxonomy. The principle of compensation expresses itself in a particularly dramatic way here. As we step up the scale of complexity from plankton to higher plants, the cost of culturability becomes higher.

---

[4]The allometric analyses discussed in this volume by Calder and Platt are relevant here.

However, the open growth system of the plant allows for high developmental and genetic plasticity. The restrictions on culturability are compensated by these modes of adaptability. As one moves to vertebrate species (such as fish) the growth system closes. Genetic and developmental plasticity are much more constrained than in the plants. Either the niche must narrow, or new compensating modes of adaptability must develop. This of course corresponds to the high state of development of behavioral and immunological plasticity in the vertebrates. Clearly, the role of patchiness changes in the higher taxa. Benthic patchiness involving higher plants or vertebrate patchiness in the form of schools of fish still allow for a high degree of independence of units, hence serve to prevent local disturbance from having a global impact. But culturability is no longer a desirable form of adaptability in these systems. It is the morphological and behavioral plasticities which serve to absorb and dissipate the disturbance. Furthermore, the high degree of internal organization in the individual organism and the social organization of the patch serve a whole variety of specific foodgetting and reproductive adaptations which must be distinguished from adaptability.[5] This internal organization also provides powerful mechanisms of information processing which increase adaptability by improving anticipation. Improved anticipation increases adaptability by allowing the repertoire of possible behaviors to be used more effectively, rather than by increasing the size of this repertoire.

Further Remarks

Unfortunately our *gedanken* device of the gene engineer may not be so farfetched. It is likely that in the future governments or enterprises will attempt to increase the productivity of the seas by using the techniques made available by recombinant DNA technology. It might be preferable to avoid this line of development. But if we cannot, the results will inevitably be disastrous, unless proper account of adaptability is taken.

The important point to recognize is that it is the repertoire of possible behaviors of a system which enable it to absorb and dissipate disturbances. The repertoire may represent a fixed potentiality, as in the fixed repertoire of possible patterns of gene activation of a bacterium. Or it may represent a more open, less preordained potentiality, as is the case with the repertoire of possible gene structures which can emerge in evolution. The repertoire of possible acceptable modes of behavior of a dynamical system is also a potentiality. But by and large, purely dynamical descriptions patterned on the dynamical equations of physics are inadequate for dealing with the rich evolutionary and behavioral dispositions of biological systems. Dynamical descriptions can hardly be expected to circumscribe the genetic and developmental phenomena which lead to evolutionary novelty. Nor can they capture the computational processes which generate the very rich behavioral repertoire of the vertebrates.

Adaptability theory incorporates the concept of potentiality into its formal structure. The quantity $H(\hat{\omega})$ is a measure of potentiality. It has to be admitted that at this stage very little is known about the dynamics of potentiality in ecological and evolutionary systems. But it is possible to pose significant questions in a conceptually coherent way. How, for example, do allometric variations in organism properties within a given taxa affect the costs of different forms of adaptability, and how do these affect the patterns of marine organization? How does the frequency of environmental disturbance bear on the adaptability structure of a marine community and on the rise and fall of adaptability during

[5]Recall that adaptability is the ability to adapt, whereas an adaptation is a system property conceived of in terms of its functional or evolutionary significance. All adaptability is a form of adaptation, but not all adaptation is adaptability. Of course, the term adaptation is sometimes used to refer to the process of adapting.

succession? Biological oceanography provides an ideal arena for addressing such questions, first because of the great range of marine environments, and second because much work has already been done on identifying and quantitatively comparing the relevant physical parameters. By correlating these physical studies with studies on the spectrum of community adaptability, biological oceanography may succeed in making a truly fundamental contribution to biological and ecological thinking.

References

CONRAD, M. 1972a. Statistical and hierarchical aspects of biological organization, p. 189–222. *In* C. H. Waddington [ed.] Towards a theoretical biology. Vol. 4. Edinburgh University Press, Edinburgh, Scotland.

1972b. Stability of foodwebs and its relation to species diversity. J. Theoret. Biol. 34: 325–335.

1975. Analyzing ecosystem adaptability. Math. Biosci. 27: 213–230.

1976a. Patterns of control in ecosystems, p. 431–456. *In* B. C. Patten [ed.] Systems analysis and simulation in ecology. Vol. 4. Academic Press, New York, NY.

1976b. Biological adaptability: the statistical state model. Bioscience 26: 319–324.

1983. Adaptability: the significance of variability from molecule to ecosystem. Plenum Press, New York, NY.

LEGENDRE, L., AND S. DEMERS. 1984. Towards dynamic biological oceanography and limnology. Can. J. Fish Aquat. Sci. 41: 2–19.

MACARTHUR, R. H. 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36: 533–536.

MARGALEF, R. 1973. Some critical remarks on the usual approaches to ecological modeling. Invest. Pesq. 37(3): 621–640.

ODUM, E. P. 1953. Fundamentals of ecology. 3rd ed. W. B. Saunders, Philadelphia, PA.

ULANOWICZ, R. E. 1980. An hypothesis on the development of natural communities. J. Theoret. Biol. 85: 223–245.

Information Theory and Self-Organization in Ecosystems

H. ATLAN

Department of Medical Biophysics, Hadassah University Hospital, P.O. Box 499, Jerusalem, Israel

Information theory, like statistical mechanics, helps to describe in a global probabilistic way, not so much what we known about systems, but rather that which we do not know. In particular, I agree completely with Dr. Rosen that complexity must be distinguished from complication. The former is a negative feature of the knowledge we have about something. I once defined complexity as an apparent disorder in systems where we have reasons to believe that an order exists. However, this order is hidden from us, and can only be approximated.

The more one knows about the details of the dynamic structure of a system, the less one needs to use probabilistic methods to describe the behavior of this system.

Even very complicated behaviors possessing self-organizing properties can be described by complicated dynamics for which many attractors can be found and where random perturbations can cause the system to go in an unpredicted way from one attractor to another. We shall see a brief example of such self-organizing activity in the appendix.

However, when everything is not known, especially when one is unclear as to how one level of organization is going to interfere with another (more integrated or less integrated, in multilevel hierarchical organizations as living systems), information theory at least can be of use to clarify some concepts, if not to provide tools for making predictions.

I have been using information theory, not so much to solve problems of stability (although I will try to say something about that) as to approach the question of how self-organization may be accompanied by an increase in complexity. I approached the issue on the evolutionary scale and also at the level of ontogenesis, where self-organization can be used as an alternative (or at least a complement) to the metaphor of the genetic program as it has been transposed from computer sciences to developmental biology. This approach also may be of some use in ecology, inasmuch as ecological systems can be viewed as self-organizing.

I will try to answer the following questions: How can complexity grow out of a less complex system? What does this emergence mean for the system itself and for the observer? Addressing this question will lead us quite naturally to the related one of how the creation of information or functional meaning might be possible.

I shall also consider the widely debated issue of stability versus complexity in terms of a dynamics of self-organization where an increase in complexity takes place at the expense of a reduction in redundancy (or mutual information at a given level of organization). I will conclude with a report on recent work which uses the formalism of random boolean networks to produce a model of a self-organizing network wherein we can identify something akin to the creation of meaning.

Within the framework of information theory, self-organization can be described as a process by which random perturbations or noise acting on the channels of communication in an organized system are able to produce, not only disfunction and disorganization, but also a change in organization to a state with more complexity and less redundancy. This is the so-called "complexity from noise" principle, which I have been using as the basis for a formal theory of self-organization (Atlan, 1968, 1974, 1981, 1983).

The point I now would like to stress is that the basic premise used to formalize this theory of self-organization by the positive effects of noise, relates to a change in the level of observation. To review it here briefly, the effect of noise acting on a channel of communication between x and y is to reduce the information transmitted through the channel by an amount called the ambiguity, as counted negatively. For example, one of the Shannon formulae for the information transmitted from x to y is,

(1) $T(x; y) = H(y) - H(y/x)$,

where $H(y)$ is the information content or complexity of the output y, and $H(y/x)$ is the ambiguity-function, produced by the noise. However, if one is interested in looking, not at the level of the channel output but at the scale of the whole system (containing x and y as related parts), then the same ambiguity function $H(y/x)$ will be counted positively, since

(2) $H(x, y) = H(x) + H(y/x)$.

In other words, the contribution of x and y to the overall information content, or complexity, of the whole system increases as the ambiguity between x and y increases. Thus, the sign of the ambiguity changes, depending upon the level at which the observation is made. Moreover, this change in the level of observation reflects a change in the level of the organization itself. For example, in a cell, where protein synthesis acts as a channel of communication between DNA and proteins, that which is seen as detrimental by the cell (errors in protein synthesis) may be seen as beneficial (i.e., an increase in variety and diversity of cells) at the level of the organism.

These reflections mean that we must expand our discussions to consider the transmission of information from one level to the other as it occurs in a multi-level organization; and that this cross-level information must not be regarded only in the limited sense of Shannon, but in its more complete sense, that is, attention must also be payed to meaning, defined here as the functional effect of information on the receiver.

Thus, the fact that a multilevel system is able to utilize random perturbations means that it is able to create new meanings for some of the information transmitted from one level to another. This creation of new meaning is what reorganization is all about. We, as external observers, do not have access to these meanings, since we see only the end products—the changes in the structure and performance of the whole system. By neglecting meaning in making use of the Shannon formalism I was able to state the complexity from noise principle with the appearance of a paradox (noise creating organization). In fact, it worked as a double negation: the destruction of a kind of information for which the meaning is absent is one way for us, as external observers, to describe the creation of meaning within a natural system in those cases when we do not have enough control over the system to directly observe the information transmitted from one level to the other.

Returning to the formal representation of this idea, let us consider the relationship between the information content, H, of a system and its redundancy. Redundancy is defined as $R = (H_{max} - H)/H_{max}$, or else by $H = H_{max}(1 - R)$, where H_{max} is the maximum information content of the system if there were no communication between the parts, and R (the redundancy) is the relative reduction in H_{max} due to the constraints between the parts. In the most general sense redundancy is the measure of what knowledge about one part can tell us about the other parts. It is closely related to mutual information. [In a two component system, the mutual information is nothing else than $T(x; y)$, which can also be written: $T(x; y) = H(x) + H(y) - H(x, y)$ or $(H_{max} - H)$, since $H(x) + H(y) =$ the maximum information content disregarding connections and $H(x, y)$ is the actual content accounting for connections between x and y.]

Now, the two different effects of noise, the one negative, the other positive, manifested at two different levels, appear in a natural way when we take the derivative of H with respect to time, where time is taken to represent the cumulative effect of all possible noise-producing or perturbing factors:

$$(3)\quad \frac{dH}{dt} = H_{max}\left(\frac{-dR}{dt}\right) + (1 - R)\frac{dH_{max}}{dt}.$$

The first term, where R decreases with time, expresses the fact that the constraints between the parts, their mutual information, is destroyed by the noise. As a result the complexity of the system increases.

But what is important is the functional complexity, i.e., what is relevant or has a possible meaning for the system itself according to what it is doing. Not every perturbation which decreases R and increases H will be meaningful and appear as an opportunity for the system to reorganize itself. Thus, the second term, where H_{max} decreases with time, represents the usual detrimental effect of noise and accounts for the fact that some perturbations are going to produce disfunctions that cannot be reutilized, and thereby will result in the destruction of relevant complexity. H_{max} may be looked upon as the output of a channel going from the system to the observer through which only the relevant information is transmitted, i.e., only that information which is functional for the system in the sense that it gives a possible meaning to what the ensemble is doing.

One does not have to postulate that this meaning is known to the observer (and this sufficiency is what justifies the use of probabilistic information theory, where the actual meaning of the messages is not taken into account). It is enough to postulate that such a meaning exists, in much the same way that an implicit but unknown source of information is assumed to exist whenever we use the H function as a measure of the information content of a system under observation. This point can be clarified by seeing what happens when we observe transmission of information at two different levels of organization (see Atlan, 1981, 1983).

One of the results of the analysis following from equation (3) is the description of the organization of a system in terms of the dynamics of its changes in complexity or information content under the effects of random perturbations. Organization, even at a single level of integration, cannot be measured by a "degree" expressed as a single number. It is always a compromise between maximum redundancy and maximum diversity (or uncertainty) and must be represented by at least three variables: H for functional diversity, R for redundancy, and t_M for reliability, a factor expressing the inertia of the system response to noise. M. Conrad in developing his theory of adaptability discussed how, according to Ashby's law of requisite variety, biological diversity may be viewed as a response to environmental uncertainty.

Self-organization appears when up to a certain dose of noise, the first term in the equation dominates the second one, so that H increases in time. But this cannot continue forever, because this increase takes place at the expense of redundancy, and redundancy must eventually be exhausted (unless additional processes can occur. We shall come back to this point later.)

This discussion allows us to define necessary conditions for self-organization to occur:
— first, enough initial redundancy must exist because it is used as a reservoir, or potential for self-organization;
— second, the system must possess enough inertia, i.e., its reliability or resilience must suffice to keep small perturbations from immediately destroying it. This is assured whenever t_M is an appreciable fraction of the interval of observation.

Now, information theory has been applied to ecology mainly by identifying the H

function as a measure of the complexity of an ecosystem. In fact, it is an index of both the diversity and the unevennness inherent in the species distribution. However, complexity is defined elsewhere as the product of the number of different species times the connectivity and interaction strength between the various species.

This latter definition is opposite to the former, because connectivity is one measure of mutual information between the parts. As we have seen, mutual information is more related to redundancy and serves to reduce the overall diversity. However, if the H function is used to measure the diversity of flows rather than of species then the case of maximum connectivity between the species yields zero mutual information between flows! Hence, let us stick to the usual measure of complexity and see how self-organization properties appear at the interface between different levels of organization where it is possible to define different kinds of average mutual information.

As Ulanowicz (1980) puts it, "Communication [between the parts of a system] can occur through a multitude of mechanisms, but to an observer who can perceive only flows, communication is limited to flow pathways". Thus, one level of organization is described by Rutledge et al. (1976) by computing an average mutual information for ecosystems based on the fractions of total throughput flowing from one compartment at time t_1 to another at time t_2. This is a kind of information transmitted between the parts, where the parts include the subsequent states of the system. When this mutual information is scaled by the total system throughput, it is what Ulanowicz calls the community ascendency. This ascendency takes into account the dynamics of the system and its functional features only to the extent that more efficient flow pathways have larger weights than less efficient ones.

However, even to an observer who can perceive only flows, the behavior of the system and its functional organization will appear to be determined by mechanisms for communication and internal constraints between the parts that are different from mere matter and energy flows. For example, structural constraints in space, such as signalled territorial limits, or a given behaviour (let us say a migration of a species) functioning as a signal for another behavior (the migration of another species), may work in the system as another kind of communication giving rise to direct and almost instantaneous flows of information (often within the same trophic level, in contradistinction to flows of matter and energy which in general go from one trophic level to the next). These additional flows of information could be used to calculate another kind of average mutual information, which would then describe the organization of the system at a different level of observation.

Behavior patterns, such as attack, escape and menace already have been considered as normal ways of animal communication (Legender and Legendre 1983). The role of indirect flow pathways may be also viewed as the (functionally instantaneous) transmission of information at a different level of organization. Obviously, these indirect pathways can transmit information between several predators (or several preys), at the same trophic level, interconnected as they are via common preys (predators). And Patten (this volume) has shown that the relative influence of indirect pathways increases tremendously when storages can occur. These storages take the guise of self-directed flows which introduce delays and amplify the number of potential indirect pathways.

These two different levels of organization and of observation could correspond to two different time scales. Over the longer scale the relevant index is the average mutual information between states at times t_1 and t_2 and would correspond to the level of H_{max}. That is, H_{max} would be an average mutual information seen by an observer of the flows during the interval from t_1 to t_2, or else an average mutual information from the system at t_1 to the same system at t_2.

This same index may also be interpreted as representing functional diversity to the observer. It is measured by considering the different flows with their relative efficiency

as in Rutledge et al. (1976). In general, random perturbations or noise acting on the system can only decrease this functional diversity, because the disturbances destroy flow pathways that are actually functional and make the system less diversified.

But this H_{max} is a maximal, because it does not take into account additional internal constraints by which one compartment is affected by the other, in a virtually instantaneous way; these constraints arising out of direct information flows between the parts tell us something about a given compartment knowing the state of another compartment (possibly in the same trophic level) at the same time. These constraints effectively act as a redundancy (or a mutual information between parts, reducing the actual diversity, or complexity, H, of the system as compared to its maximum H_{max}). Random perturbations usually act on these constraints to make them less stringent; in other words, they tend to reduce this redundancy, or mutual information at the level of instantaneous observation, and thereby increase the complexity of the system.

More generally, what I have called constraints by direct instantaneous transfer of information between parts refers to constraints observable within short time intervals; and what I have called constraints between states at t_1 and t_2 means constraints observable over a much larger time scale. Now, one can go a step further and postulate a third level at which the interplay itself between these two levels might be modified. Whereas redundancy, being normalized, cannot be higher than 1, mutual information can increase ad infinitum, providing the increase in H_{max} is unrestricted (see Ulanowicz 1979, 1980). However, this rise in H_{max} cannot happen simply by the addition of different parts. In order for it to be functional—to have a meaning—this increase must come from new interactions at a different level of organization. A preliminary increase in the redundancy of the system at a different level of organization, e.g., by non-useful multiplication of repetitive material, or by the multiplication of possible functionally equivalent pathways may serve as a potential for the eventual creation of new functional diversity (Milgram and Atlan 1983).

In the process of self-organization that I have just described, redundancy can only decrease, and this decrease is means by which there is an increase in specificity and complexity. However, as in other instances of self-organization, such as non-directed learning, one has to postulate an additional mechanism by which redundancy can be generated in order to start the self-organizing process; or can be regenerated after the initial redundancy has been depleted, in order that the process may continue.

In the scheme represented by equation (3), an increase in redundancy must be accompanied by a decrease in the complexity, H, of the system. Therefore, since we do not want to lose the complexity which already has been gained, the process by which redundancy is initially produced or regenerated must happen at a different level, where it is possible for the creation of R to proceed with no concomitant loss in total H. This is possible if an increase in the size of the system would take place (1) with no loss of the diversity already created; (2) no built in specificity, i.e., with a large number of functionally equivalent pathways. It is this kind of initial functional equivalence, interchangeability of multiplicity of choice which Ulanowicz (1980) calls redundancy, and it is expressed as a conditional entropy which measures the uncertainty attributable to this multiplicity of choice.

This initial lack of specificity is also what is called degeneracy and may be found in the genetic code as well as in the initial states of the repertoire of the immune system and also in the initial poly-innervation, which subsequently is reduced during the course of development of the central nervous system. More generally, in a non-directed learning process, where the patterns are produced by the elimination of elements and pathways and by the creation of more and more specificity within a class of initially equivalent patterns, such a recharging of redundancy can be produced by recreating a state where all the initial associations which progressively had become forbidden during the learning process are

once again allowed, providing that the emerging patterns had been stored in a memory so that what has been learnt is not lost. I have postulated (Atlan 1979) that in our cognitive system, where non-directed learning can be viewed as self-organization, the role of paradoxical sleep and dream, where forbidden associations are once again allowed, is to recharge our neural functional redundancy so as to keep our learning capacities going.

At the evolutionary scale, such a creation of new redundancy may have occurred by the addition of genetic material in the form of additional chromosomes or repetitive DNA that would have been initially neutral and functionally interchangeable.

In a dynamic ecosystem, where many different levels of organization are interacting with one another (different in that the levels are characterized by disparate scales of space and time) general statements about the effects of complexity or redundancy on the dynamics at one level can always be countered with examples coming from observations on the effects at a different level of organization. For instance, the classical complexity–stability problem can be restated as follows:

First, as I mentioned before, what is usually called complexity in ecological studies is in fact also a measure of redundancy, since it combines species richness with connectivity and interaction strength, i.e., with functional constraints between the species.

Then, ecosystem stability for the most part has been studied in terms of local stability, i.e., investigating the probability that a dynamic system will go back to its stationary state after it has been perturbed. In this approach one takes into account possible non-existing systems, with their probabilities of occurrence in comparison with existing states. As Pimm (1984) eloquently puts it, it "involves a comparison of existing and hypothetical systems". Such comparisons are particular examples of changes in the level of description, as can be seen clearly from the effect of unevenness in species abundances on the system diversity. Richness and unevenness are put together in the classical measure of diversity and complexity — the H function. It is well-known that this function is maximal when all the probabilities of occurrence are equal, and the Shannon formula then reduces to $H = \log N$. This measure is justified for a given system when the a priori probabilities are estimated among all possible hypothetical observations, while the existing system is a realization of only one of them.

On the contrary, actual evenness, i.e., equiprobability in the observed outcomes, implies uncertainty due to an actual multiplicity of choice among equally functional systems, i.e., a redundancy or degeneracy that is opposite to complexity or diversity. More precisely, the classical complexity–stability study by May (1973) has shown that, for a given species number, increasing the connectivity increases the probability of system instability in the linear dynamical sense. This result does not mean that a system with high connectivity cannot actually exist: although its a priori probability of occurring under the assumption of random assembly is low, its particular structure can happen to make it stable.

Now, if a stable system does not react to random perturbations in the environment (like crystals to high temperature), i.e., if its inertia, resistance to noise, functional reliability or resilience keeps it from immediately being destroyed by perturbations, it will respond to noise by decreasing its redundancy (connectivity). It thereby increases its complexity according to the complexity from noise principle described earlier, and its probability of being stable will consequently increase.

Thus, this hypothetical system would have a remarkable evolution. At the beginning, it would be stable in spite of its redundancy, such stability being the result of exceptional structural and functional conditions. However, its redundancy would serve as a potential for self-organization by helping it respond to noise by increasing its complexity. At the same time, its stability becomes less and less exceptional (in the sense of May), since its

redundancy is decreasing. Thus, it is approaching a state of maximum probability as its complexity increases. (This tendency to maximum probability is not necessarily a good thing, as it parallels the evolution of living systems towards death, or maximum entropy.)

In fact we return to the problem of the limits on using information theory to talk about complexity and organization. The difference between dead complexity and living complexity is, of course, the functional state of the system, i.e. the *meaning* of its organization in terms of possible function. It is well known that classical, probabilistic information theory does not address the meaning of information.

However, my goal in describing the theory of self-organization through the positive effects of noise has been to be able to describe the creation of new meanings in organized systems without having to know how actually to formalize the meaning of information. Again, using noise, i.e., random perturbations, to relax the constraints at one level produces some heterogeneity, and this generally contributes to disorder and misfunction because it is done randomly, i.e., with no relation to any previous or planned state of order. However, up to a certain limit, if the introduced malfunctions are not enough to kill the whole system, the very same heterogeneity can be seen as a new state of order — one having more complexity and less redundancy available for use at a higher, more integrated level. In other words, moving from one level to the next higher produces a change in the sign of the effect of noise on a channel of communication — a change from negative when it is subtracted from the information transmitted in the channel, to positive, when it is regarded as a measure of the additional variety introduced by loosening the constraints. The very fact that heterogeneity survives and keeps functioning at the higher level of organization (for example, heterogeneity of cell functions when considered at the level of the organism, heterogeneity in the individual organisms when considered at the level of the species or heterogeneity at nearly instantaneous time scales when considered over long durations) implies that it is being used in a new, organized way, which in turn means that it has been transformed into a new kind of homogeneity.

This going back and forth between levels allows us to understand how changes in organization can take place in time, in a manner that appears autonomous to the eye of an observer having only partial knowledge of the constraints between the various elements at all levels. By simultaneously observing several levels of integration in a non man-made, organized system one encounters many logical difficulties but also becomes aware of the richness possible in the analysis of such systems. It has been argued that information is necessarily transmitted between one level and the next, and the meaning of this information plays a central role in the functional organization of the whole. However, this meaning is not known to the observer, because, in general, he has empirical access only to each level by itself, and not to what passes from one stratum to another. The complexity from noise principle indirectly expresses the role of these unknown, but crucial, meanings in self-organization. Again, relaxing the constraints at one level creates not merely disorganization at that level, but also functional complexity at a different, more integrated level, where the new relations are integrated into a new functional organization having more diversity and less redundancy. By the way, that is why a high initial redundancy is a necessary, but certainly not a sufficient, condition for self-organization to occur. The system must also be able to incorporate the revised state of connections in a new functional way. So, self-organization when viewed as noise-induced disorganization followed by reorganization implies interplay between the different levels, and describing self-organization in this way amounts to describing the creation of new, but still unknown, meanings in the information transmitted from one level to another. In other words, what appears as organizational randomness to an observer outside the system implies the creation of some new meanings, yet unknown to him, within the system itself.

Now, the meaning of information in natural organizations is very different from what it is in artificial constructs, where a final goal is defined, and it is in relation to this goal that signals or connections either have a meaning or are meaningless. In a natural organization the final goal is not known, and meanings must be defined according to self-created criteria, which may appear very unexpected and far-fetched to an observer, who would have to imagine such meanings, a priori. These self-created criteria may also be transitory.

To illustrate self-created meaning I would like to propose a model of a network of interconnected elements, working as if it recognized binary sequences, where the structure to be recognized is defined by a randomly influenced substructure of the network. Thus, randomly created algorithms are shown to be able to distinguish between classes of patterns. In particular, pseudo-random sequences are recognized as non-random because they belong to a specific class which is defined by the particular structure of the network able to recognize it. But this structure itself is the result of a self-organizing process, i.e., it is partially the result of a random process. Thus, the criterion for distinguishing among sequences defines a kind of meaningfulness (or lack of meaningfulness) of the sequences and is nothing else than the algorithm of recognition itself. This algorithm cannot be defined a priori, i.e., it makes no sense to speak of a particular meaning before the network is built.

In this model (see appendix) we make use of a known self-organizing property of random boolean networks which evolve from random homogeneous initial states towards final, organized configurations, where structures in space and time can be observed as arrays of subnets of stable elements separated by subnets of oscillating ones. After such a network has reached its final (generally oscillating) state, one (or more) specific pathway between two elements can be identified in such a way that one of the elements serves as input for strings to be recognized and the other as output. The behavior of this output element defines the codes for recognition and non-recognition.

The model demonstrates a mechanism by which to divide messages into those which are recognized and those which are not, while the criterion for this demarcation, which is analogous to making sense and not making sense to a cognitive system, is nothing else than a given inner structure. The structure has no meaning other than being able to produce this demarcation, and may itself have come about randomly.

It is as if complexity, which appears as an apparent non-reducible randomness, could be removed by a kind of orderliness that did not come about as a result of planning but itself resulted from apparent indeterminacy and randomness. This, in my opinion, is the consequence of the close relationship between complexity and disorder that exists in natural systems not planned or ordered by man. The only difference between complexity and disorder seems to be the existence of an apparent meaning or function of the former in the eyes of the observer.

In conclusion, information theory has proven to be a useful tool in quantitative ecology. However, in its classical form, where the only quantities are static measures of probabilistic distributions having no functional meaning, it has severe limitations. The extension of information theory into a theory of self-organization that incorporates the interrelationships between different levels and the creation of meaning, may be particularly relevant to the analysis of evolving ecosystems.

References

ATLAN, H. 1968. Application of information theory to the study of stimulating effects of ionizing radition, thermal energy, and other environmental factors. Preliminary ideas for a theory of organization. J. Theor. Biol. 21: 45–70.

1974. On a formal definition of organ-

ization. J. Theor. Biol. 45: 295–304.

1979. Entre le Cristal et la Fumée. Seuil, Paris, France.

1981. Hierarchical self-organization in living systems, p. 185–208. *In* M. Zeleny [ed.] Autopoiesis: a theory of living organization. North Holland, NY.

1983. Information theory: basic elements and recent developments, Chapt. I, p. 9–41. *In* R. Trappl [ed.] Cybernetics: theory and applications. Hemisphere Publ. Co., New York, NY.

ATLAN, H., F. FOGELMAN-SOULIE, J. SALOMON, AND G. WEISBUCH. 1981. Random boolean networks. Cybernetics and Systems 12: 103–121.

AUSTIN, M. P., AND B. G. COOK. 1974. Ecosystem stability: a result from an abstract simulation. J. Theor. Biol. 45: 435–458.

FOGELMAN-SOULIE, F., E. GOLES-CHACC, AND G. WEISBUCH. 1982. Specific roles of the different boolean mappings in random networks. Bull. Math. Biol. 44(5): 715–730.

KAUFMANN, S. 1970. Behavior of randomly constructed genetic nets, p. 18–37. Vol. 3. *In* C. H. Waddington [ed.] Towards a theoretical biology. Edinburgh Univ. Press, Edinburgh.

LEGENDRE, L., AND P. LEGENDRE. 1983. Numerical ecology. Elsevier, New York, NY.

MAY, R. M. 1973. Stability and complexity in model ecosystems. Princeton Univ. Press.

MILGRAM, M., AND H. ATLAN. 1983. Probabilistic automata as a model for epigenesis of cellular networks. J. Theor. Biol. 103: 523–547.

PIMM, S. L. 1984. The complexity and stability of ecosystems. Nature 307: 321–326.

RUTLEDGE, R. W., B. L. BASORE, AND R. J. MULHOLLAND. 1976. Ecological stability: an information theory viewpoint. 57: 355–371.

ULANOWICZ, R. E. 1979. Diversity, stability and self organization in ecological communities. Ecologia 40: 295.

1980. A hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

Appendix: A Model for Self-Created Meaning

The process goes as follows: We start with a random network of boolean automata of the kind first studied by Stuart Kauffman (1970; see also Atlan et al. 1981). Every element in the network is an automaton which can be in either of two states. The state of each element is the result of a boolean function of two binary input variables. There are 16 possible kinds of such elements defined by the 16 boolean functions. Each element defined by one of these functions receives two inputs, one from each of its two other neighbors. In a typical experiment, a network of such elements is constructed as a 16 × 16 matrix, closed on itself in such a way that the last element in any row is connected to the first element of the same row, and the columns are closed in a similar manner. The different possible boolean functions are then randomly distributed on the elements of the network as initial conditions.

Well-established results (Atlan et al. 1981; Fogelman-Soulie et al. 1982) are that any given network exhibits an asymptotic behavior that is relatively stable to variations in initial conditions. After a relatively small number of iterations (no more than 100 or 200, run in parallel) the network enters a limit cycle with a relatively short periodicity.

In addition, a definite pattern of subnets appears in the network. It is characterized by a set of subnets wherein all the elements are oscillating and these oscillating nets are separated from one another by other subnets wherein all the elements are stable. The cycle length of the whole network is the smallest common multiple of the cycle lengths of the individual oscillating subnets. For a given network, this pattern of subnets is relatively stable to changes in the initial conditions. Some elements are always stable; some are always oscillating, and some others ("wishy-washy") are either stable or oscillating, depending on the set of initial conditions.

Figure 1 depicts an example of such pattern formation for the same network under different sets of initial conditions.

It is important to realize that this stability of pattern in the face of changing initial conditions is not really local stability in the sense that the system possesses an attractor to which it returns after it has been perturbed. It is more like what has been called buffered stability by Austin and Cook (1974), where the system moves to a new asymptotic state,

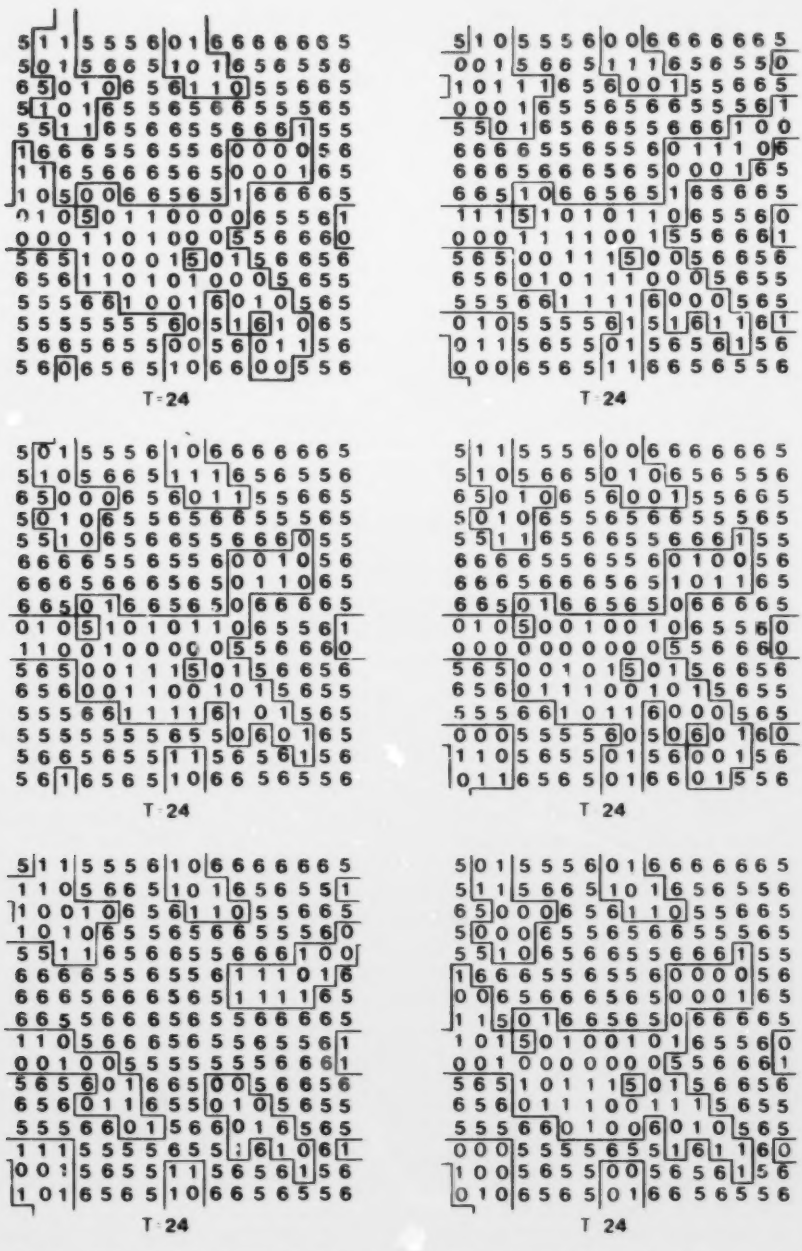
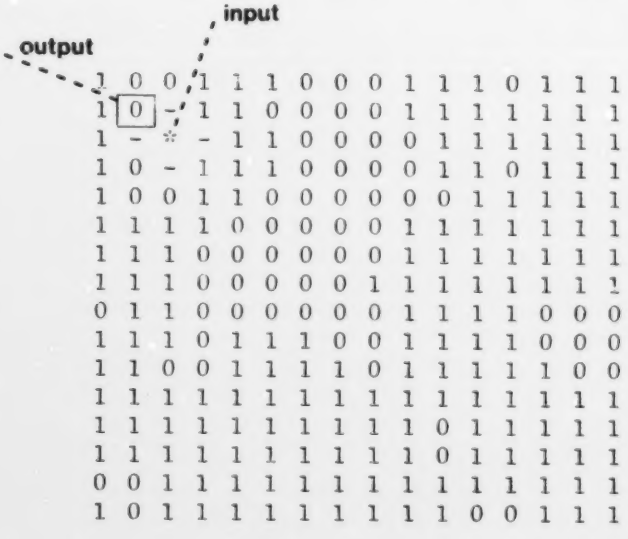


FIG. 1. Subnets. The frames display the state of the network during the limit cycle corresponding to six different initial conditions. T = cycle length in discrete time intervals. 0 and 1 indicate oscillating elements; 5 and 6, stable ones (from Atlan et al. 1981).

15	0	8	15	15	15	0	0	0	15	15	15	0	15	15	15
15	15	8	15	15	0	0	0	0	15	15	15	9	9	15	15
15	15	8	8	15	15	0	0	0	0	15	15	9	15	15	6
8	0	0	8	15	15	0	0	0	0	15	15	0	15	15	8
8	0	0	15	15	0	0	0	0	0	0	15	15	9	15	8
15	15	15	15	0	0	0	0	0	15	15	15	15	5	5	8
15	15	15	0	0	0	0	0	0	15	15	10	10	12	12	12
12	15	15	0	0	0	0	0	15	15	15	12	10	12	12	12
0	10	10	0	0	0	0	0	0	11	11	12	12	0	0	0
10	10	10	0	12	12	12	0	0	11	12	12	12	0	0	0
15	15	0	0	12	12	12	12	0	12	12	10	10	10	0	0
14	14	14	6	12	12	6	6	15	15	12	12	10	10	15	15
14	14	12	6	12	12	6	15	15	15	0	10	10	9	15	15
9	14	12	12	12	12	12	15	15	15	0	10	5	15	9	15
0	0	12	12	12	12	15	15	15	15	15	15	5	15	15	15
15	0	12	12	15	15	15	15	15	15	15	0	0	15	15	15

(a)



1	0	0	1	1	1	0	0	0	1	1	1	0	1	1	1
1	0	-	1	1	0	0	0	0	1	1	1	1	1	1	1
1	-	*	-	1	1	0	0	0	0	1	1	1	1	1	1
1	0	-	1	1	1	0	0	0	0	1	1	0	1	1	1
1	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1
1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1
1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0
1	1	1	0	1	1	1	0	0	1	1	1	1	0	0	0
1	1	0	0	1	1	1	1	0	1	1	1	1	1	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1

(b)

FIG. 2. (a) Approximations of stable (0) and oscillating (15) cores are determined by limit cycles of 15 sets of initial conditions: the number on each element indicates how many times the element was found to be oscillating in the final state; (b) Structure of the limit cycle after noise was inserted at the input element for a given set of initial conditions (0 is stable and 1, oscillating).

e.g., one with new oscillating elements or new stable ones. However, this new asymptotic state is not too far from the original one, so that the macroscopic spatial pattern of subnets remains roughly the same.

This is an example of self-organization where the change in the level of observation is crucial: at the microscopic level, once the distribution of the rules and initial conditions has been cast, the final structure is already defined. However, to an observer who does not know this detailed structure, the initial state appears as a random homogeneous distribution, whereas the final state appears organized both in time and in space. As a matter of fact, in this case, even if an observer knew the details of the initial structure, it would not help him very much to predict the end product, unless he ran the computer-simulation, which would amount to performing an experiment on natural system to determine the outcome.

Now, on one element of such a network after it has reached its final organized state, we inject "noise" in the form of strings of pseudo random sequences of 0 and 1. This perturbing string is inserted for at least two cycle lengths, during which time the new asymptotic behavior of the network is compared with that of the original noiseless network.

Figure 2a shows the behavior of a particular network for 15 different sets of initial conditions. The numbers from 0 to 15 indicate how many times a given element was found to be oscillating in the asymptotic state. Thus, the zeros indicate the stable core, the 15's represent the oscillating elements and the remainders are wishy washy. Then in the network in one of these asymptotic states noise was inserted in the element marked by a star. Disregarding this element and its first neighbours, the rest of the network showed the asymptotic behavior depicted in Fig. 2b, where zeros represent stable elements and 1's the oscillating ones. In other instances, the effect of adding noise was to force some elements of the stable core in the noiseless network (i.e., the zeros) to become oscillatory. However, it may happen, as can be seen in the example of Fig. 2b, that an element that was always oscillating in the noiseless network (2a) appears to be stabilized by the "noise".

In fact, if we try to understand the details of how the perturbing string affects the subsequent states of the network, we can identify certain features of the pseudo-random sequences which produce this stabilization. We can interpret this phenomenon as a kind of recognition process: the element where the noise is inserted serves as input to receive the sequences to be tested; when another element that is normally oscillating becomes stabilized by a given sequence, it means that the sequence has been "recognized", and this element serves as output to the recognition device.

A detailed analysis of this particular case will help to understand what is going on. It is enough to consider the elements connected from the input element to the output and to search the history of their values calculated from their boolean functions and their connections. This study is represented in Fig. 3, where the input element in Fig. 2b (where the noise was inserted) is called A, and the output element (i.e., the one stabilized in Fig. 2b) is called D. Connected elements B, C, and E are responsible for the state of D. We have to look in detail at the subsequent states of all these elments in the limit cycle, and at how signals are transmitted from A to D due to the rules governing the elements and the connections between them.

Then one can see that if 0 and 1 are seemingly randomly inserted on A from the outside, the element D, which is normally oscillating in the noiseless network, becomes stabilized in state 1 under the conditions that when C is in state 1, A is in state 0. Otherwise, when C is in state 0, D remains in state 1 regardless of the state of A and this indifference is what is important for our analysis. It is due entirely to the oscillatory behavior of element C, to the stable behavior of E (in state 0), and to the rules which compute the signals at B

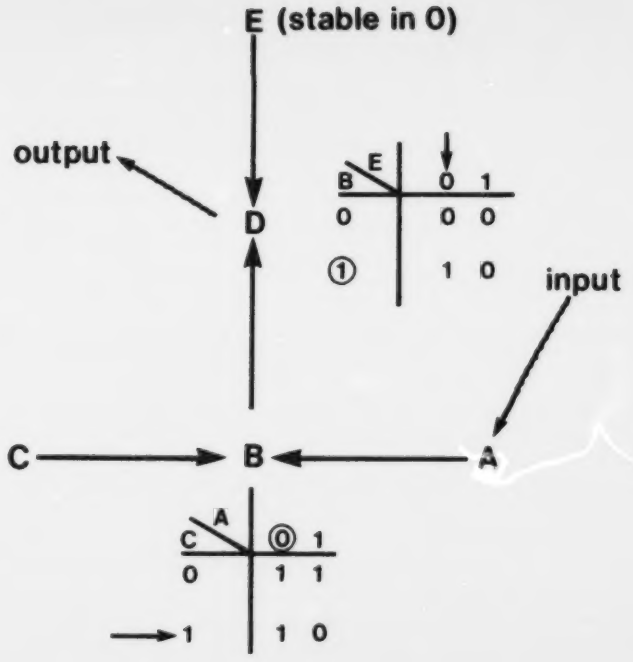


FIG. 3. Detail of the connections affecting input and output elements. C is oscillating with cycle length 8 repeating the following sequence of states: 01010110. D is stabilized in 1. Since E is always in 0, if D is to be in 1, B must be in 1. In order for B to be in 1, it is sufficient that when C is in 1, A be in 0. When C is in 0, A is inconsequential. As a result, in the input sequence, a 0 in A is necessary to match a 1 in C: *0*0*00. The other input digits (*) are indifferent.

and D as represented in Fig. 3. C is oscillating with a very short period of 8 time intervals, and it goes through the sequence 01010110 repetitively. For an input sequence to stabilize D it is sufficient that its 2nd, 4th, 6th, and 7th digits should be 0 in order to match the 1's in the sequence of C and cause D to stay in state 1. The other 4 digits in this sequence of 8 are inconsequential.

The overall period of the whole network in the example was 96, which means that the inputs were sequences of 192 digits, half of which were arbitrary with respect to this recognition problem. Therefore, it is possible to build a long sequence by repeating any 8 digits, so long as the entries 2, 4, 6, 7 are constrained to be zero. Since the others entries are arbitrary, there is a large class of pseudo random sequences with this hidden periodic structure in whose members will be recognized by our device. Applying any pseudo random sequence to the input element A, one can expect that it will be recognized if, and only if, it belongs to this class.

Thus, a specific pathway in the stabilized network from the input element to an output defines a class of sequences that will be recognized.

From Hydrodynamic Processes to Structure (Information) and from Information to Process

RAMON MARGALEF

Department of Ecology, University of Barcelona, Barcelona, Spain

Introduction

What we know about oceans and ocean life is based on the study of samples, often obtained without much prior planning. Innocent hypotheses about the statistical properties of the pelagic universe are made and inferences drawn concerning local variability and changes with time. Technical developments and new sampling procedures produce, with increasing frequency, transects of point data that place the whole problem in a different light. It is not wise to accept these different measures as disconnected and mix them in the same bag, to be used with the most simple statistical hypotheses: the structure of the data has to be preserved carefully.

Perhaps we should take for a reference transects obtained from aboard a ship as the only objects with which to analyze spatial structures in detail. Two dimensional pictures from the air or from space do not yet allow enough detail. Except when crossing dynamic fronts, where transitions can be relatively sharp, the values of each variable along the transect constitute a "mountain range" profile, or brownian sequence, in which deviations are random relative to the values in their neighborhoods, with a degree of contagion between points that can be proportional to the rate of increase of the populations and to their diffusion. The diversity, as a variance, can be computed not only at each point sample, but also over a more or less long string of pooled samples. The spectra of diversity centered on successive points may change gradually or abruptly, in general, "flat" spectra of diversity are associated with "mountain range" patterns in the values of the most important variables, and diagonal spectra are associated with distributions of several variables in which contagion is not so evident. Moreover, the transects or sequences change their character, sometimes gradually, sometimes in a discontinuous way, and more frequently than not, changes in one direction are not symmetrical with those in the opposite direction.

The ordinary transects, as they are analyzed, postulate a heterogeneous distribution, but do not imply more regularity than is usually expressed by the word "patchiness." With the present state-of-the-art it is possible to make hypotheses only about the distribution of different variables in space and time that are compatible with the observed transects. For two or three dimensions, the data base is too poor to check the suitability of existing mathematical methods.

The conception of the organization and of the function of the pelagic ecosystem is evolving and should keep pace with the interpretation of the details shown in transects. The model used to describe the processes that generate the ecosystem (conceived as a quasi-steady state) needs to be improved. In the traditional approach, distributions were believed to be approximately uniform, with a certain random patchiness that could be fitted by some simple statistical distribution or explained by simple differences in the diffusivity of the various interacting elements. A supply of nutrients from outside could be added to local recycling, perhaps forced into the system by external or exosomatic energy. The most common models assumed boxes or compartments with given rates of transfer. Eventually, fluctuations, or else a steady state, could be achieved through a "clockwork" model.

Now, a more realistic approach is to keep in mind that no machine turns twice in exactly

the same way, much less an ecosystem in which energy is being continuously converted into organization and information. In this paper I insist on the necessity of replacing such earlier models by other more complicated ones, in which a number of events, or instabilities, are generated with time, through inputs of external energy at more or less disjoined centers and with different intensities. Each one of them starts an ecological succession that, usually, can be associated with the disappearance of the vertical turbulence and the development of a certain measure of stratification. Such a model would be more consistent with thermodynamics and system theory. The average of a number of such processes going on in the watermass, at different speeds, with different phases, and with different probabilities of resetting to the initial conditions, might well produce average values comparable to those used in more simplistic models. My feeling is that we need the more complicated model for at least two reasons: (1) to accommodate the actual observations along the transects, and (2) to replace a uniform forcing function (external energy) by a spectrum of discrete events.

The whole subject of organization can be addressed by information theory and from two points of view: how the pelagic machine converts energy into information, as reflected in the resulting structures; and how the transects can be used to reconstruct dimensional structures, or to check if the three dimensional structures generated by some pertinent hypothesis are validated by the observational transects.

The system is governed, punctuated or reset by inputs that come from outside and which remain unpredictable from inside the system. The situation is almost desperate, in that what is predictable is not important, whereas what is important is not predictable from within the system. Furthermore, we feel obliged to enlarge the system of reference until it becomes unmanageable.

From Energy to Structure and Information

The fundamental constraints on the behavior of systems are thermodynamics. Thermodynamics includes the rules of "constructivity" of nature. Ecosystems are, in the first place, physical systems. On the macro scale, when details are lost, the rules of constructivity remain.

Energy is not destroyed, but can never be observed twice in the same way. The decay of energy is translated into an increase of information, in a very broad sense, or into the generation of structure. Energy that has done work, or is of low quality is not lost from living systems, since it has allowed organization to increase, and this occurred in a particularly efficient way, because of the self-multiplicative property of information. According to the mood of the moment, we can weep for the decay of energy that eventually should lead to a cold and uniform universe, or rejoice in the incessant enrichment of structure, that could lead to an equally dreadful end: the universe as a frozen "objet d'art." I hope that both ends are equally far away, and that ordinary things happen in between. To explore this middle ground within the realm of biology, and especially plankton biology, is the purpose of this paper. Most things have already been said elsewhere, but I hope to combine them in another prespective. We need to take into account the discontinuity of living matter (individuals), and to distinguish two levels in energy: internal or endosomatic energy, necessarily of "high quality" (electromagnetic, chemical bond), and external or exosomatic energy, thermal and mechanic, in the environment around the organisms.

A Basic Model for Phytoplankton Dynamics

The general properties of nature and of ecosystems are represented in a model that has been proposed for phytoplankton production (Margalef 1978b):

P	=	I	+	(A	×	C)
Production		Inputs (advection)		Turbulence as external energy		Covariance in distribution of factors of production, or functional interconnection

in a more abstract form:

Non-renewable resources	Renewable resources	
	Energy momentum tensors	× Divergence operators

The advection (of nutrients) is also dependent on external energy to transport them from outside to inside the system of reference. Note that external energy is reflected not only in production, but also in organization.

Volterra–Lotka models have various disadvantages: they ignore the discontinuity of individuals (quantification) and space, and operate in a sort of "ether" of very unreal properties. Taking into account that the covariance (mass action law) of distributions is multiplied by a coefficient of probability of interaction (a_{ij}) analogous to a turbulence, and the existence of inputs not predictable from inside, the equivalent of the former expression is, in terms of the Volterra–Lotka formalism,

$$dB/dt = \sum_{i=1} R_i + \sum_{j=1} a_{ij} N_i N_j$$

stochastic	probability of interaction governed by covariance of distributions, and specificity

Without inputs and outputs, former formula reduces to $P = A \cdot C$, and, if we take the derivative with respect to time:

$$d^2B/dt^2 = dP/dt = [a.\ (dA/dt) + b.\ (dC/dt)]$$

acceleration or deceleration of production (B, biomass)	decay of exosomatic energy	segregation of elements or increase of information

The last expression is satisfactory for several reasons. It includes accelerations and decelerations, that is, forces. It is more meaningful to start from this last equation, rather than from Volterra–Lotka expressions, to explore the connections with thermodynamics. It is easy to see that

$$d^2B/dt^2 - b.\ (dC/dt) = a.\ (dA/dt)$$

growth (as a force)	− differentiation (eventually, diversity)	= forcing energy

in good agreement with experience. The former expression describes succession, best characterized by the decrease of the ratio P/B (production/biomass), or defined from "less mature" to "more mature" states (Margalef 1968).

The same expression also summarizes other ecological knowledge. To keep a system exploitable we can assure an input of non-renewable resources from outside (dimensions of force) that do a definite amount of work; but it also helps to increase the mixing inside, also due to exosomatic energy, and prevent further organization of the system. Exosomatic

energy makes the interaction less dependent on topological and historical constraints; but as energy decays, spatial organization of the system develops.

When the frame of reference is enlarged, the former inputs of nonrenewable resources become internalized. Any distinction between advection and turbulence, any estimate of the turbulent energy, or any measure of the covariance in the distributions of the potential reactants (light, nutrients, cells), are obviously related to the size of the cell of observation. A spectral expression is necessary throughout.

Our purpose is to use the observed distribution of individuals and species as an evolving message, and from there to go back and deduce previous situations, trying to visualize the processes that stand behind the quality, intensity, and regularities of the observed pattern. In consequence, the most interesting part of the general expression concerns the arrangement of the different interactive components of the ecosystem in the form of covariance, or its inverse, the segration. I refer to the covariance in the distributions of different cells, nutrients, etc.; when variance is low, the reactants can be assumed to be segregated, but this happens at different scales. For instance, the concentration of chlorophyll and the concentration of nitrate may be negatively correlated, or segregated, over short distances, but their respective averages may be positively correlated over long distances. When the segregation within a vertical column is considered, light behaves singularly in that its extinction in water always keeps down the covariance in the vertical dimension. Light cannot be pushed mechanically, and the covariance in the distribution of different factors of production can never be strictly proportional to mixing and turbulence; even if organisms move up and down, their concentrations will never be positively correlated with the distribution of light. If energy is present in the form of turbulent diffusion, one has to relate it to the size and motility of the individuals of the different species, making the application of this model more complicated, although not fundamentally different.

Systems

The essence of the ecosystem is the pattern that links its components. The a priori freedom of behavior of each of the parts is more or less diminished when the parts join the system. But the elements retain some flexibility or elasticity, because of time delays at the junctions. Ecosystems behave in just this way, being made up of individuals of a certain size, of behavior more or less unpredictable, and out of equilibrium most of the time.

The ecosystem is not a rigid machine made of gears and levers, but involves more costly transmission of information across a fluid (and turbulent) environment. In one sense it is comparable to a Turing machine, an automaton able to read information from a tape, and to use it for any purpose, including writing it unto a new tape, or for building other machines. However, it makes little sense to keep separate the instructions to build the machines, to operate them, or to process and pass information to other machines. The blurring of proper distinctions between operative parts and memories makes organisms and ecosystems quite different from computers, and this has to be kept in mind in formulating ecosystem models. They cannot be rigid clockwork.

The ecosystem retains some elasticity or flexibility; otherwise we should think of it not as a system, but as an object. Th slogan "everything is connected to everything else" may be used to good purpose by ecological activists, but is rather bad pedagogy. Patten (1982) has suggested an intriguing aspect of ecological networks in proposing that indirect relations among the components of an ecosystem may count more than direct interactions. I feel uneasy before any notion that explicitly or otherwise accepts that each component is rigidly or deterministically linked to everything else. I prefer to see some internal

flexibility as one of the essential properties of any system that merit the name. The comparison between a string of cars in a railroad train, and a platoon of automobiles along a highway may express the differences that I have in mind. Probably the only really relevant connections in the ecosystem are those involving negative feedback loops, and those that are the basis for the acquisition of information.

The elements of the system are replicable units, that is, species, each represented by a number of individuals. The eventual degree of indeterminism, as well as the capacity for "decision" at each node, seems to be related to the ability to accumulate information locally. As individuals, living beings seem able to accumulate information in direct proportion to their individual biomass, and in inverse relation to their turnover, that is, in proportion to B^2/P.

We should refrain from excessively simplifying any conception of the ecosystem to the point of being only a net of interactions, especially if some analogy with physics is to be drawn. Individuals are unique systems by themselves, much different from the elementary particles of physics. Individuals are the subjects of natural selection, and not merely the carriers of atoms of certain elements, or of chemical energy. As a consequence, the relations between the level of the ecosystem, as a channel, and the level of its components (the individuals) as the bearers of messages, may be analogous to the relations between a discourse and its constituent words, with comparable reciprocal influences pursuant to any change.

Many systems, not only ecosystems, share the property of being composed of parts replicable by themselves (organisms) or by an extrinsic agency (viruses, words, electronic or mechanical parts, etc.). The parts are available in any number (words) or are produced in excess numbers; mutual dependence and natural selection adjust the relative numbers of the parts. New parts can be generated, and are accepted or rejected by the workings of the other components of the ecosystem. Some very fundamental components will remain, almost forever with small change; more sophisticated components, however, will often need the prior development of other elements for structural and functional support. Accepting that the components are systems on their own, and able to evolve (species, as well as words, or mechanical parts), one conclusion that I deem very important, because it is fully applicable to the global ecosystem, is that the overlying structure of a system slows down rather than enhances the possible rate of evolution of each component. The concept of system assumes that each element, after joining the system, finds its freedom limited. Genotypes can evolve faster than phenotypes, as the study of neutral characters has shown. Nevertheless, free from some of the constraints of the ecosystems, e.g., under domestication, plants and animals have shown an increased capacity for change. A certain stability in the contingent of words or of their meaning is essential in a language. Shifts in the numerical proportions between pairs of species are more rapid in disconnected parts of ecosystems than in whole communities.

In the framework of the ecosystem, evolution has led to an early split between primary producers and consumers. The transport of energy as chemical bonds has encouraged a further differentiation of the animals into the multiple links of a food chain. Relations between size and turnover become pertinent. It is easy to visualize the relations between plankton size, turnover and trophic position. Information accumulates at different levels: genetic, ecological, and cultural. Genetic information is a relatively well defined concept, encompassing the levels DNA, RNA, enzymes and substrates. The first agent acts more as a specifying channel, the last is more labile. Cultural information reflects individual history in the form of personal memories and of artifacts. Human artifacts and trees are magnificent examples of material expressions of cultural information. Ecological information consists in the present species, in their relative numbers, in how they are spatially

arranged, and which net of interactions they support. Information can shift from one channel to another, but the total organization is dependent for its accretion, even for its conservation, on energy being degraded in the system.

Now there are a number of rules or regularities concerning the way information appears to increase at different levels, and the most profound question concerns the apparent non-coincidence of the areas of energy decay with the regions of information gain. Some of the elements that contribute to this property of systems are obvious: a large computer is more powerful than two simple half-sized computers. Extant information acts as a selective channel that helps to accumulate more information. There is also a subtle shift in the quality of information. Interactions between predator and prey are paradigmatic of the diversifying property of the whole system. Such interactions engender different statistical distributions for prey and predator. As the predator lives longer, it interacts again and again. The prey evolves characteristics of immediate utility for survival, whereas the predator tends to improve its learning ability. The oceanography of the present day appears unable to reconcile the view of individuals as carriers of matter and energy, with their rôle as subjects of natural selection.

Considered as a whole, it is clear that the ecosystem receives an input of electromagnetic energy at one end of its organization, comprised of the primary producers; but the relative importance of the accumulators of information at all the levels increases towards the larger predators. The energy from turbulence influences the elements of the ecosystem in a different way, according to their size. In general, for the whole system, decay of energy is superimposed on an inverted gradient of accumulation of information. This is a reflection in the structure of the ecosystem of the operation of what has been named the principle of St. Mark (4:25), "For he that hath, to him shall be given: and he that hath not, from him shall be taken even that which he hath". The ecological interpretation is that the result of interaction among ecosystems or parts of ecosystems is such that information increases more in the system or subsystem that was already more complex to begin with.

These considerations have direct implications in modelling. One can feel sympathy for the hypothesis of unstructured food webs (Isaacs 1972; Lange and Hurley 1975) as a reaction against elementary deterministic constructions, but food webs cannot grow on a neutral terrain, and the whole system is or becomes anisotropic in entropy and information. One need remember only that the fraction of energy going into respiration is not the same across the food web. If the net of relations is to be cut into simple binary systems of predator and prey for the purpose of analysis, then the non-coincidence of the energy decay and of the increase of information becomes a source of difficulties. Binary elementary systems are probably not sufficient, and ternary systems should be considered. I have discussed this elsewhere (Margalef 1980).

On Varieties of History

Information can express itself as oscillations (e.g., internal rhythms) or wax and wane with the passage of time. Wavelike changes in an ecosystem cannot be symmetrical. Symmetrical waves are possible only without inertia, and in a world of pure energy. In ecological models we speak of the relaxation times of the different components, and their combination and interference in systems of many elements, if one has to trust simulations, generally results in a dampening of the amplitude of oscillations, and a lengthening of the periods. This agrees with observations of the natural world. Although complicated systems are not devoid of fluctuations, the most violent change is usually encountered in systems with a rather low number of component species.

The non-necessarily recurrent change, involving shifts in information content, has the following basic property: organized change or ecological succession is slow, the result of information accreting through the extant structures, but such successions are punctuated by catastrophic change. The biosphere has the appearance of a quilt made of a mosaic of heterogeneous patches, since the changes one way (successional) are slow, and the changes the other way (disasters) are fast, it appears that the whole surface of the earth is living evidence of succession. The concept of standard succession has been much critisized, but is reasonably well-documented (Margalef 1974). It is interrupted and set back by catastrophes or regressions that appear random and are independent of the density and other properties of the populations (Acevedo 1981; Hanson and Tuckwell 1981; Margalef 1980). Extreme disasters (e.g., impacts of planetoids) are less frequent than climatic fluctuations, and the short-period cosmic rhythms (yearly, daily) have been assimilated by organisms as a matter of course. Now an important question is whether there is some inverse relationship between the magnitude of a disturbance and its average time of recurrence.

The intensity of disturbances is inversely related to their frequency, and the spectral distribution is quite often comparable to the cascading of energy through decreasing wavelengths of turbulence. Probably it is unwise to speak of regular cycles, concerning either El Niño's, or other climatic fluctuations. The spatial distribution of energy perturbations, in principle, may be random, and comparable to their distribution in time, but in some places the spectrum of disturbance can be systematically extended, in the sense that large catastrophic change is more common in such places than elsewhere. The slope of the spectrum, however, need not be different. Catastrophic events can use energy released from inside the same system, as in a fire in the forest, or in the use of fossil fuel or of nuclear energy by man (cf. Atlan, this volume). Disturbances accelerates change, and erases memory or information at different levels. The later stages of succession develop more internal constraints and slow the turnover. The system uses information to impede the entry of more information, thus retarding further development.

The development of phytoplankton in a lake, as turbulence decays and stratification develops and increases, goes from an initial peak in productivity, with a prevalence of diatoms and small flagellates, towards an increasing presence of dinoflagellates and later, perhaps, of bluegreens. The nature of these latter stages depend on the time available; the system is frequently reset to an initial condition through disturbances, in the form of wind, that mix the water. The seasonal overturn has the quality of a major and, perhaps, more regular disturbance. Zooplankton has to be less sensitive than phytoplankton to the smallest disturbances; herbivore lifespans are longer than those of phytoplankton cells. Migration and, in general, movement of the zooplankton overcomes in part both temporal and spatial perturbations. We should be prepared to recognize different frequencies and modalities of environmental change as expressed selectively by the different layers of the organization of the ecosystem (cf. Conrad, this volume). Summarizing: the late stages of an ecological community depend on the time available for carrying out the process of organization before the whole system is reset by some disturbance, and the degree of the disturbance defines the initial stages and tempo of the new successions.

The asymmetry of change (slow succession, fast and catastrophic regression) appears also in abiotic historical systems, as in the development and destruction of thermoclines, and in the cycles of sedimentation. In sediments, a gradual transition is common from coarse to fine materials, followed discontinuously by a new event of coarse sedimentation. In such situations, the causal dependence on available external energy is obvious, and has the same sign as the historical changes in ecosystems.

Elsewhere (Margalef 1980) I have refered to the two modalities of change as "history

making" and "history telling". From the point of view of thermodynamics we can recognize, respectively, a rapid degradation of available energy, and a behaviour more in accordance with that of open systems far from equilibrium, in which there is a trend to decrease the amount of energy exchanged per unit of preserved structure.

This slowing down of historical change is related in some ways to our understanding stability through expressions associated with the name of Liapunov (1892). Let N_i be the density of each of the species present then

$$\sum_{i=1} (dN_i/dt) \to 0, \qquad \sum_{i=1} (d^2N_i/dt^2) \to 0$$

This convergence that could define stability, if such a concept is worth defining in ecology, can coexist with an indefinite divergence at a very small scale, a scale in which reactivity is limited and the energy available for changes is very small. This divergence may be expressed in terms of the differences between numbers in every pair of species,

$$\sum_{i,j=1} d(N_i - N_j)^2/dt\,.$$

All these expressions can be written with reference to time, or to space, using t or x. In relation to space they should express a trend towards a general uniformity that, however, does not go to total uniformity. Alternative spatial configurations should always be able to persist below a certain scale.

There are many different kinds of disturbances and the responses of ecosystems are also diverse. In situations such as upwelling areas, or in eutrophic systems, both in freshwater and in the sea, the ecosystem is fed from outside in that it receives an input of nutrients, or of organic matter, but it would be hard to say that our reference system is exploiting the neighboring systems, because the control remains outside. It may be convenient to use a special designation for such a relation; I propose to qualify the force fed system as stressed or subjected to stress. The response of systems such as eutrophic lakes and upwelling areas is twofold: acceleration of turnover (increase of production) and the exclusion from one fraction of the cycling elements, either as an internal passive storage or, more frequently, at the boundaries (oxygen and nitrogen from the water to the atmosphere, phosphate and organic carbon to the soil or the sediment). This was introduced in an earlier expression:

$$P = I + (A \times C)$$

input (e.g., eutrophication) } forms of stress

(e.g., upwelling) } forms of stress

increased production — segregation at the boundaries } response

This concept of stress reminds us of the importance of buffering mechanisms and of couplings between open adjacent structures.

Measures and Messages

The structures of an ecosystem, as revealed along a transect, can be viewed as a message. Although we should not forget the reason the Tiger-lily told Alice, "We can talk when there's anybody worth talking to": in the case of plankton it is probably wise to abandon any hope of finding much detailed meaning in the strings of values along a

transect. Probably, we might expect to identify a certain "style", nothing more. This style has a certain generative process, represented by differential growth of the different kinds of symbols (point values), and of transpositions, shifts or changes of place. In other words, in the same way that the observation of the frequency of association between letters and spaces allows us to mimic languages, and to imitate the styles of particular writers, it is hoped that a careful analysis of the combinations of symbols could provide inferences about the processes that are locally responsible for the apparently garbled messages that are detected along the transects.

The pattern may be an expression of many mechanisms and operations conceivable as processes. Such processes are obvious in a cloud, and in a plankton patch, its scale being such that individuals are seen shifting against an unstable background, or subject to a sort of thermal agitation — in contrast with the image of a forest or of a benthic community, where a two dimensional pattern is more persistent and noticeable, because it approaches the rigidity of position of the letters in this page. Conditions of competition among fixed organisms are different from conditions in plankton, so also with respect to the length of life. The study of transects is a first approach, that has to be complemented as other possibilities become available.

The appropriate measures would concern the relative amounts and the interspersion of the different components, for instance, of the different species. Measurements have to be chosen so as to obtain similar numbers for systems with equivalent properties. Probably, it is impossible to get scale free or dimensionless measures and space and absolute size always count.

Measurable structures are the result of processes that can be expressed by their rates. If the rates are relatively unconstrained, as in one organism that grows, or a species that is filling a space, or even an ecosystem that redevelops after a catastrophic event, different measures $A, B, C \ldots$ of variables like number of cells, total volume of cells, concentration of chlorophyll, etc., may be found to approximate relations of the type,

$$k(dA/dt) = l(dB/dt) = m(dC/dt) = \ldots,$$

or else the proportionality might be established among the second derivatives.

This means simply that if the proportionality among the differential rates is maintained, the development takes place with an external manifestation of regularity. Imagine an organism or an elongated part of an organism, that grows faster in length than in thickness: as it grows, it becomes more elongated. Or think of an assemblage of phytoplankton, in which the small-celled elements grow faster than the larger cells. The ratio small cells/larger cells changes accordingly. $A, B, C \ldots$ may be continuous or meristic. The number of organisms is usually taken as continuous, although it is not, and this becomes important in the case of low densities, for large individuals, and also at the level of the introduction and extinction of species in the ecosystem.

When $A, B, C \ldots$ are the numbers of individuals of different species, part of the emergent regularities are discussed under the heading of diversity.

$A, B, C, \ldots$ may be measures of different properties of one population, such as volume, dry weight, chlorophyll, or ATP in phytoplankton. Large cells are, in general, more hydrated, contain less chlorophyll and divide more slowly. Local and instantaneous values have to be estimated against the background of such relations. It is hopeless to trust linear conversion factors. Bacteriologists are now going through the same agony that phytoplanktonologists have known, in trying to evaluate bacterial biomass through measures of ATP, or of muramic acid.

Metric allometry is better known than chemical allometry. It has been applied to the study of growth, and to the changed proportions among related species of different size.

Still, many taxonomists stubbornly believe that differences in ratios are more important than differences in absolute size.

The usual expression takes the form $A = aB^k$; logarithmic transforms of the original measures are appropriate, and, of course, sets of measures $A, B, C, \ldots$ relative to the same sample cannot be all normally distributed. Discontinuous or meristic characters (vertebra, scales, number of teeth in combs, number of elementary structure in diatoms) follow the same expression with k usually close to 2/3. This means that a space of one-half size is neither filled with half the number of structures, nor by the same number of smaller structures, but by a number somewhere in between. In two dimenions, as in cell mosaics or in diatom structures, the number of structural elements grows approximately as the 1.3 power of linear dimensions.

To speak of fractional powers introduces us to the realm of "fractals" so enticingly presented by Mandelbrot. Fractals are also found in the total number of species in relation to the total number of individuals in a given space (or with the measure of this space), in the number of realized connections in relation to the number of possible connections; in the number of branchings and the length of the segments in relation to the covered space in trees and fluvial systems; and in the extension of interfaces through folding at the boundaries of cells and of patches and "pancakes" in the biosphere.

This comes closer to issues more related to our main subject, like diversity and connectivity. Also to the following sort of relations: the ratio zooplankton biomass/phytoplankton biomass usually decreases as the absolute concentration of phytoplankton increases; the ratio diatoms/dinoflagellates is related to total phytoplankton; the pigment ratio D_{430}/D_{665} is very much dependent on total amount of chlorophyll (inversely related); and the diversity is related to the biomass and rate of increase of biomass. Quite often ecologists believe that the ratios mean more than the absolute value of one of the compared magnitudes, and in so doing behave as an ichthyologist who feels that to state that the diameter of the eye is 3.5 or 4 times the distance between the eye and the rand of the operculum, is more serious than to state plainly the size of the individual fish.

The fractal relation is recognizable too in the accepted dependence of energy on the wavelength measures of turbulence. All the interesting concepts with which we have to deal must have a spectral expression. Allometric and fractal considerations help to simplify the expressions of these spectra.

Diversity

A message at the level of the ecosystem consists in distributions, over the space, of individuals, that can be sorted into different subsets, usually into different species. In systems subjected to thermal or eddy dispersion, as are the plankton, we can choose to declare any instantaneous distribution irrelevant in comparison with the more persistent spatial configurations present in the benthos and in forests. Then we cannot go further, in practice, than give the relative numerical representations of the different species. If the distribution appears regular, that is, if the numbers of individuals of the different species ranked in order of decreasing abundance, suggest the possibility of replacing the whole set of data by a simple number, this number is the diversity. Diversity can be computed on any distribution of a set into subsets, but it has meaning, if any, only in relation to the criteria followed in making the distribution — size, individual biomass, chemical components, specific identity, etc. Computation of biotic diversity requires identification of the individuals as to species, and can be related to the dynamics of complex populations. Other criteria of classification — biomass in species, total volume of particulate matter in

classes according to individual size of particules, etc. — allow the computation of diversities that may have a meaning in relation with the stated criteria.

The abundance of the different species presented as a series of abundances from the most to the least numerous, may be more informative than a simple diversity, but to use a single number is justified whenever such series is more or less regular, in a system that has persisted for a while without too much external disturbance.

Diversity is necessarily spectral. At a single point, diversity is zero, and it increases with the sampled space. It is easy to prove that the same numerical values for diversity can be obtained in samples of the same size, but belonging to ecosystems of different history and structure (Margalef 1969, 1978). Intuition suggests that diversity may convey much information about the "character of an ecosystem", but it is no panacea. Low diversities, or a flat spectrum of diversity, are to be expected in pioneer communities and in turbulent environments, both of which have a predominance of fast growing opportunistic species. Higher diversities or more step diversity spectra are expected in more "mature" and organized ecosystems, where not much energy is available for further reorganization, and many species persist while represented by small numbers of individuals. A similar approach can be applied to a partial analysis of language and of artistic creations (music, paintings). Again, low diversities characterize languages that are perceived as warm by many cultures, with a high proportion of expletives; the equivalent in other areas could be a repetitive music, or forms of pop art. High diversity, or diagonal diversity spectra, associate with cold and pedantic language, symphonies and complex pictures, with a hierarchy of themes.

A careful matching of the distribution of species against the distribution of environmental factors discloses that the coincidence of a set of species is compatible with a limited set of properties of the environment. Species are very specific sensors. Phytoplankton has the advantage of reacting fast and of not surviving for a long time when conditions are no longer appropriate. The richness of organization, the ecological segregation implied in diversity, can be inversely related to the covariance of the factors of production, and diversity can be expected to increase as a result of the decay of the mechanical energy through the system.

The most common expression for diversity is borrowed from Shannon and Weaver, and has the form $-\Sigma\, p_i \log_2 p_i$, where the p_i's are the separate proportions of the species present, and $\Sigma\, p_i = 1$. The use of joint probabilities of occurrence for combinations of two, three or more individuals, introduces many possible ramifications. The equivalent expression,

$$(1/N) \log_2 (N!/N_a!N_b!N_o! \ldots),$$

where the N's stand for the actual numbers of individuals in the set, can also be convenient. For instance, diversity decreases if one individual of a rare species is taken out, and replaced by another individual of an already more common species (the usual action of man). The same expression is useful as well to decompose the community (and its diversity) into building blocks, formed by groups of species, each of which can show a different pattern of change.

In the preceding formulation, diversity is a measure of the entropy, expressing the richness of combinations that are possible in the assemblage of individuals that are present over a certain space. It can be related, especially if spectra are included, to the achievable degree of ecological segregation and organization in general. The value of the Shannon–Weaver index, in bits per cell, is usually between 1 and 2.5 in phytoplankton of coastal waters, being especially low in estuarine, polluted, and upwelling areas. Values from 3.5 to 4.5 are most frequently measured in oceanic phytoplankton, but local prolif-

erations of selected species are the cause of low values of diversity. Although diversity is usually low in the centers of upwelling areas, horizontal mixing leads to rather high diversities around, and also close to the fronts. In most oceanic areas of low productivity, a great number of species are present, with low and rather uniform population densities, resulting in diversities close to 5. But, even in quite oligotrophic areas, temporal or local development of one or a few species results in lower diversities. The presence of large numbers of small organisms such as cyanobacteria and flagellates is understood to depress diversity, although there is much uncertainty related to taxonomy. As for the slope of the spectra, in an example that is, at least, indicative (Margalef 1969), the diversity of populations taken along a stretch of one mile in well-mixed water with a predominance of diatoms, passes from 2.9 to 3.2 in point samples, to 2.9−3.4 in pooled samples; in more stratified places, rich in dinoflagellates, the values were 1.5−2.2 and 3.1−3.9, respectively.

Many other indices of diversity have been proposed. One of the simplest is (number of species — 1)/(log of number of individuals). Simpson uses as an (inverse) diversity the probability that two individuals taken out of one sample belong to the same species. The probability of extraction is considered to be independent of size, activity or any other specific property. Another good start would be to consider the probability that the two individuals extracted belong to different species, divided by the probability that they belong to the same species,

$$\sum_{\substack{i=1 \\ i \neq j}} p_i p_j \Big/ \sum_{i=1} p_i^2 = \left(1 - \sum_{i=1} p^2\right) \Big/ \sum_{i=1} p_i^2 .$$

This is a good index of diversity, equivalent to a transformation of Simpson's index that places it into a plausible scale for the naturalist. Its logarithm, in the usual range, correlates well with the Shannon index.

One advantage of the last index is that it can be used to generate other indices, as may be convenient. For instance, instead of using the probability of picking individuals of different species, we can prefer to speak of choosing two different species which are able to interact. By so doing, we abandon the notion of diversity as the description of a static situation, and go back to the forces behind, to the processes of interaction that are reflected in, the present diversity.

Connectivity

Diversity may be a measure of the capacity of the ecological channel to carry information. All combinations are allowed. The last expression includes all sorts of interspecific interactions and measures their pooled probability against the aggregate of intraspecific interaction. This expression may be especially interesting because it portrays diversity as an upper limit to real connectivity. This concept might be more realistically expressed by,

$$\Sigma\, a_{ij} p_i p_j / \Sigma\, b_i p_i^2,$$

where the a_{ij}'s are the probabilities of interaction between each pair of species (i, j), and, correspondingly, b_i is the probability of interaction between the individuals of one species i.

This may come close to the comparison between the possible reactive and non-reactive behaviour of individuals inside the ecosystem

$$\Sigma\, a_{ij} p_i p_j / (1 - \Sigma\, a_{ij} p_i p_j)$$

If a_{ij} is a probability of interaction, the possible values of a are not only 0 and 1, but may take on fractional values, and perhaps even be above 1 should one want to express some "facilitation". Dependence or interaction between species is peculiar to each kind of interaction, and is related to space in a form more complicated than simple mass action. In fact, the structure of interactions define an (ecological) space. One vexing question concerns the definition of "interaction". As the analysis of possible feedback among three species shows, the three feedback loops that can link the species, pose certain constraints. As an example, if two predators (competitors) prey on one species, we have two corresponding negative feedback loops. Among the competitors the feedback loop is positive, disruptive, and probably going to vanish. This is the way, or one of the ways, that some of the links may be cut, and connectivity thereby decreases. But it is not clear which possible links should be counted and which should not. Perhaps only the interactions that should be counted are those in which there is some exchange of both materials and information. In which case, many or perhaps all indirect interactions, like competition, should be excluded.

In 1968 (Margalef 1968), I pointed out, based on empirical evidence, that species which interact feebly with others, do so with many species. Conversely, species with strong interactions are often part of a system of few species. Mathematical simulation confirms the expected decay of large and highly interconnected systems. Common sense leads to a recognition of the limitations of connectivity as a guarantee of functionality and survival. In speech, the connection of words is not total or arbitrary, it follows a grammar. A totally connected electronic system is shortcircuited and useless. The key to the relative flexibility and small indetermination of any system is its partial connectivity. In ecosystems, if total diversity is high, relative connectivity is usually small. Diversity has an upper limit around 5.5 bits per individual; it is doubtful that this number is highly significant; but probably the range of real connectivity is more important for the preservation of active organizations, perhaps with a very general or universal meaning. (McNaughton 1978; Margalef and Gutiérrez 1983; Rejmánek et al. 1983; Ulanowicz 1980). Too few links in a dynamic system can result in chaotic behavior; too many links in rigidity. A restricted connectivity allows an increase in the number of species. Many species appearing in small numbers in relatively old and well organized systems are often dependent on special resources or on one or few species and do not contribute much to connectivity (they are not automatically connected with all the existing elements), but they nonetheless contribute to diversity. The intermittence of the interactions among individuals of different species is prevalent (except for permanently united parasites and symbionts) and probably serves as a limitation on connectivity, whatever its effects on "stability" might be. In the same sense, concerning intraspecific interactions, the distinction of the sexes may also be important in stabilizing relations among individuals.

Observations on the development of ecosystems provide examples of a limitation on connectivity as the number of species increases. Further introduced species are usually typified by a few individuals with complex life styles and a strong dependence on only a few of the extant species. This is the regular pattern in the increase of the number of species in an ecosystem which goes undisturbed for a long time. In the seasonal development of ecosystems, the relations become more specific and segregated, as for instance in the relations between flowers and insects. Habituation and learning processes lead as well to a decrease in relative connectivity.

As these properties seem to be general to all systems, relations observed in alien but comparable fields may be suggestive. Electronic circuits are helpful because they graphically illustrate connections and the integration of subsystems. Integrated circuits provide a partial paradigm for the evolution of eukaryota from prokaryotic elements. We might

consider whether primitive forms of life, without a centralized mechanism of transmission, could be compared to ecosystems rather than to present eukaryota. Another interesting comparison is offered by construction toys such as Meccano or Erector. Components are used to build mechanical models, that must undergo a selection (like an ecosystem); can be taken apart; and the same parts (elements or individuals, belonging to a set of classes, or species) can be used again, in different proportions, to construct other models. The distribution of the total number of parts into the different kinds of parts is perfectly analogous to the distribution of individuals into species, and values of diversity can be computed, with the same meaning. Diversities around 5, close to the maximum, characterize richly articulate and functional models. Nuts and bolts are, of course, the most abundant components, comparable to bacteria or phytoplankton. Not all the parts are connected to all the other parts, and the actual network of connections is clearly a subset of the maximal connectivity identifiable with high diversity. Theoretically, if possible connections are closely proportional to the square of the number of elements; connections in real networks are fewer and come closer to the power 1.6, again fractal quantities.

Language and neural nets, in addition to electronic circuits and mechanical models, pose problems of their own albeit with many analogies to ecosystems. Useful as such analogies are, they cannot describe all the problems that are encountered in the study of ecosystems. Inbuilt thermodynamic asymmetries in food webs, transport systems, etc., might provide supplementary rules of constructivity.

From Information Back to the Generating Processes

Inputs of external energy generate dynamic processes in the ecosystem that are perceived as starts or resettings of successions. These inputs irregular in space and time, impose small or large retrogressions in the degree of organization. When seen from inside the ecosystem of reference, such steps backward occur at random, however, the process of ecological succession at each place may nevertheless follow predictable regularities. If this is true, patchiness has to include components that are not totally random, for instance, the regular evolution (succession) of the patches. Also, there must exist a certain hierarchy in the global structure: small structures inside large structures, organized according to a pattern that, at least in terms of scale, has to depend on the regularity of the spectrum of energy inputs. From the biological point of view, the observable results are the distributions of species, their relative abundances and their interspersions. From the observation of present distributions it may be possible to infer or reconstruct in part the past history of the system and the processes that have been active therein.

Boundaries

Systems do not blend with other systems in a gradual way. We recognize distinct galaxies. Segments of discourse are bounded within the covers of each book. In the pelagic environment, both the cascading of energy, and the existence of a bottom and irregular coastal topography, are causes of much heterogeneity. The usual distinction of macroscale (oceanic gyres, upwelling systems), mesoscale (fronts, eddies, rings), and microscale (down to the level of molecular diffusivity) reflects the reality of a certain degree of discontinuity, of boxes within boxes. This discontinuity is present also in relation to time. The differences are as between a smooth, thin liquid stream and a dripping faucet.

Thinking in terms of pieces of machinery, there is a relation between the size of a machine and its capacity to use energy of a certain quality to perform work. Organisms, being miniaturized, need high grade energy (short wave light, chemical bonds). Thermal

energy in wind and waves is effective only over a certain spectral "window". Gradients in separate properties can combine scissorlike to produce rather sharp boundaries, as in the thermoclines, in chlorophyll maxima between light and nutrients, in zonation of freshwater macrophytes due to the interaction of light and hydrostatic pressure.

This is not the place to discuss the hydrographic and biological heterogeneity present at all scales in the oceans, but I do wish to point out a substantial difference between vertical and horizontal heterogeneity. The stratification related to the axis defined by light and gravity is sharp and is basic to the organization of ecosystems. Many hydrographic phenomena contribute distinct packages of mixed water of definite properties, but in the end they spread, as boluses and pancakes, according to their density. Their volume may be well mixed internally, but they are well separated from overlaying and underlaying strata by small but sharp differences in density across which there is little diffusion. With the exception of Langmuir cells, etc., horizontal differences are usually less sharp and not as persistent, and this lack of contrast leads us to treat horizontal differences over small scales as stochastic. The distribution and the properties of phytoplankton are much dependent on the distribution of turbulence, and the small scale physical structure of the environment becomes a framework to which the "plankton machine" complies.

There are different sorts of boundaries. Horizontal boundaries between a surface water layer and a deeper layer are an especially active part of the organization. It is instructive to perform mental experiments and guess the consequences of placing walls in different positions in the sea: an horizontal wall would lead to strong processes of reorganization in each of the separate compartments. Each would tend to reorganize along the axis defined by light and gravity, showing a polarity comparable to that displayed in the morphogenesis and regeneration of organisms.

Imposed vertical boundaries would be, on the average, of less consequence, but would cover a wide range of possibilities. They can be of high tension and rather sharp (the *limes convergens* of Dutch writers (van Leeuwen 1963)), as in the mesoscale structures (fronts) or those cases where the system on one side of the boundary exploits that on the other. These high tension (In using this term I am drawing an analogy with surface tension.) boundaries may appear to migrate in response to hydrographic events and also as a result of ecological succession.

Other vertical boundaries can be described as low tension, being more imprecise and corresponding to the *limes divergens* of Dutch plant ecologists. They occur between much less dynamic ecosystems, that may be different in taxonomic composition, but are usually much less diversified in terms of biomass and productivity. Such boundaries are obviously fractal, tend to be evanescent, develop nebulous topographies, are often very complicated, and seem comparable to patterns in the chemical composition of clays and humic substances or to the geometry of rivers. Such low energy, highly complicated structures (dead ends in a certain way) are also found in evolution. From the point of view of thermodynamics it is easy to compare all these terminal situations with much of what has been variously described as the "salt of the Earth" or the "baroque of Nature".

Coarse Grain Patterns

A pattern is some synoptic configuration which contains some repetitive motifs and allows an abridged description. In this sense it is appropriate to speak of pattern as the form or style, as in literary or musical composition. The patterns in ecosystems are expressions of history, and usually much stochastic detail is added as they are being generated. The study of ecological pattern is motivated by their descriptive (and sometimes their aesthetic) value, and also by the wish to identify the processes engendering them. In plankton small

scale turbulence and a whole spectrum of hydrographic discontinuities over larger scales, define the minimum dimension over which one may fruitfully study pattern. Analysis is feasible at the scale of 10^1 to 10^3 m with interesting results. Patterns observable in this range can be refered to as "coarse grain patterns".

The pelagic ecosystem can be considered as being made up of an array of columns or vertical fibers, each one organized around the axis defined by light and gravity. Each column supports a complete ecological cycle, with excess production near the top, and excess respiration below. Probably the following expression is the most concise description of the driving forces in an unstressed column that lead to a vertical segregation of the factors of production:

$$\frac{\int_0^{-z} \text{primary production } dz}{\int_0^{-z} \text{total biomass } dz} \rightarrow \text{minimum},$$

where the depth $-z$ may be either the maximum depth of the photic zone or the total depth of the water column. The principle applies to both situations, but the second configuration offers a better view of the role of heterotrophs. Primary production is a result of sufficient covariance among the distribution of the factors of production.

Each vertical column may be considered as an oscillator. At the very least there exist the day/night rhythms in primary production. These, and other oscillations, propagate horizontally through the adjacent columns. Perhaps a more careful analysis of biological oscillations would be justified. The expressions of Volterra and Lotka might be applied, not to species, but to larger blocks of components of the ecosystem, such as, in this application, to the upper layer containing the primary producers and other prey, and to the deep layer with its predators.

With many simplifications, the rate of growth (r) of the population in the top (T) is $r = A \cdot C/T$, and the "mortality" (m) of the population in the deep layer is $m = V(dD/dz)/D$, where A is energy of turbulence, C is covariance in the distribution of reactants, T is the biomass of the population in the top layer, V is a fictitious sinking speed that also includes true mortality, and D is the biomass of population in the deep layer. From the respective relaxation times of each plankton community we could expect biotic oscillations of frequency

$$F = \frac{1}{2\pi}(r \cdot m)^{1/2} = \frac{1}{2\pi}\left(\frac{A \cdot C \cdot V}{T \cdot D}(dD/dz)\right)^{1/2}$$

This, at least, is not counter-intuitive, and it may be helpful to explore conditions near the depth at which the most intense daily rhythm in the properties of the populations is to be found. Should waves exist with a larger amplitude or longer period than the daily rhythm of production, they would introduce more complications into the basic structures induced by the more or less discontinuous inputs of external energy.

The whole system is driven by impacts of energy from outside. In a first approach to representing the behavior of the system, such impacts may be considered more or less pointlike and discontinuous, distributed randomly in intensity over space and time. This is not completely correct, because such events can be correlated or have a contagious distribution by reason of the nature of the atmosphere and hydrosphere, and besides, energy cascades through a sequence of definite structures. In principle, impacts of external energy — from raindrops to large meteorites — reset ecological succession by mixing, and these effects propagate spatially according to the local values of turbulence introduced or

already present. Perturbations are not necessarily pointlike or linear, e.g., fronts, but it is expected, from hydrodynamic considerations, that when no important stress or shear is present, that they end up as more or less circular eddies.

The inputs of external energy combine functionally adjacent columns, with the result that each one of them is no longer a relatively independent system. This convolution is very important in that it affects stability, and it can profitably be compared with terrestrial ecosystems: a tropical rain forest in the plain can be modelled as a bundle of independent columns (a "small is beautiful" model with negligible horizontal transport); whereas a rural + urban complex, or a river, cannot be understood without realizing that much energy is invested in horizontal transport ("big is powerful"). Stability should always be discussed with reference to the horizontal dimensions of the stabilizing mechanisms.

It is only natural that the availability of mixing energy results in a functional complementarity between adjacent columns that are no longer alike. The more turbulent spots become centers of production and sources for the diffusion of the produced materials. The horizontal structure of the ecosystem near the surface becomes patchy. At depth, horizontal differences are much less. Vertical migration of the animal components of the ecosystem is very effective in damping incipient differentiation in the deep layers.

A Volterra–Lotka approach has been proposed for explaining the evolution of biotic waves that expand horizontally. The same approach to considering the adjacent populations in the top layers predicts the appearance of asymmetries in the topological relations between patches characterized by the dominance of different components. The degree of asymmetry depends on the rates of growth and diffusion (Levin and Segel 1976; Steele 1973). Spots with a predominance of primary producers should tend to form discrete patches embedded in a honeycomb of other patches in which other organisms, eventually consumers, are more common. This pattern can be replicated at different scales, and reinforces or is associated with the discontinuous quality of the inputs of energy. In other words above a minimum size below which patchiness may be random, the size of the structures is governed by the rates of growth and the speeds of diffusion. At still larger scales the distribution of the energy impacts modulates patch size. The same basic principles of horizontal organization are found in the humanized landscape (Margalef 1979), although in a more persistent form, and are probably a common feature of all ecosystems.

These considerations are supported by only casual observations, never sufficiently detailed, and by spectra computed on transects of continuous measurements of physical and biotic variables. It is easy to envision that the patterns should be changing continuously, contracting, expanding, intensifying or vanishing. By analogy, one might think of a structure drawn on a piece of rubber that is continuously being deformed. In relatively empty water, a large structure can grow at an accelerated speed, thereby accommodating an increase in internal differentiation. Perhaps some comparison can be drawn between this latter case and morphogenesis in embryos.

Inversion of the topological relations among the patterns also may be worthy of study, as when Langmuir cells change from stripes to nets, or when spots become stars and converge into honeycombs. Structural changes may be particularly important in mesoscale phenomena (fronts, etc.) and in structures containing strong vertical boundaries (dinoflagellates in the stratified water on one side of the front; coccolithophorids confined to the front, etc.).

It is, then, necessary to add a vertical dimension to our models. Models, such as those used in cosmological speculation (Centrella and Melott 1983), could provide some valuable suggestions, as they describe the contrast between expanding structures that are rather rounded, and zones of convergence that often have the appearance of collapsing and extended filaments linked in a sort of reticulum. This particular model suggests that it is

more appropriate to speak in terms of pancakes than of patches when referring to certain basic elements of structure. But the three dimensional models used by cosmologists are not totally applicable to the realm of pelagic ecosystems, where the vertical dimension has a unique quality.

In summary, the marine environment should be considered as a space with certain properties, in which a number of disturbances are created at random, propagate and die. There is probably never a lack of seed, and populations of organisms wax and wane in accommodation to local properties and local change. Properties of life eventually slow down and impart an historical character to the accumulation of information in these "downhill" processes. The biosphere has developed the ways and means to cope with fluctuations and impacts not exceeding a certain amplitude, and this capacity provides a common background for interpreting much local change. In this scenario, deep waters probably exert a stabilizing role.

Study of Transects

The general picture traced in the preceding section is based on many observations, summarizes many facts, but remains an hypothesis. Observations from the air or from space can provide some data on patchiness, and especially on the sharper boundaries. Probably the best way towards testing and eventual validation of my hypotheses is through the study of transects; the characteristics observed along transects have to be compatible with the three dimensional distributions postulated.

The transect turns out to be a message: a spatial series of physical properties, fluorescence, particles, taxonomic composition or sonar records. A sufficiently detailed taxonomic study may be considered as a string of symbols, and there are many possibilities for analysing such sequences: Markovian chains, frequency spectra and, in general, the resources of information theory pertaining to the study of messages. Perhaps a very effective procedure might be to calculate a sequence of diversity spectra, each one of them centered at a particular point, and enlarge the sample in the direction of the transect. Notice that I propose using diversity at three levels: point diversities, spectra of diversity, and local differences between the diversity spectra.

The usual way to study transects is to record local intensities, and to use these data to calculate average values, and to detect trends, gradients and periodicities. Spectral analyses can be performed, for instance, in an attempt to compare biological distributions to spectra in physical properties that in certain regimes is related primarily to turbulence. Patterns are never symmetrical: peaks may be eccentric, not only in relation to, for instance, the fetch, but also to the dynamics and migration of populations. Taking a transect requires appreciable time, and therefore information is not synchronic. The use of Doppler effect to explore patterns taken from two ships running along a track in opposite directions should be undertaken. Sonic tomography may be helpful in describing the structure of pancakes.

The statistical approach to the study of plankton has been crude. It does not make much sense to mix samples in a bag and pretend that the distribution resulting therefrom says something important. Data like the abundance of species, when plotted along a transect, produce a "mountain range" profile with fractal quality. Each local value is related to the neighboring points in a way that more or less can be considered random. Thus, the departures from the averages resemble a Brownian pattern. In the usual statistical distributions, the variances are much higher than the means, and are approximately proportional to the mean raised to a power between 1.6 and 2. Logarithmic transformation is used to normalize the distributions, and things usually are left at this point.

In the typical pattern of plankton patches high density loci are surrounded by a net or honeycomb of lower density regions; the image of mountain ranges along the transects is appropriate. After logarithmic transformation, the profiles become more symmetrical, losing the appearance of peaks. Often one can write,

$$\log N_{x+1} = C \log N_x + R,$$

where R is a normally distributed random number, and C is empirically related both to the rate of increase of the species, and to the coefficient of diffusivity. The problem is complex because the density at point $x + 1$ is related not only to density at point x, but also to the densities over all the neighboring points.

In the hypothetical situation where there are empty areas in the water, the expected profiles would be different, and a symmetrical or normal profile could be obtained with the transformation $\log (K - N)$, K being a constant upper limit. No pattern of this sort is observed for phytoplankton distributions, but the model can be appropriate for animals that move around and for the netlike pattern around the centers of disturbance and of production. In such situations the following expression will probably apply:

$$\log (N_{x+1} - N_x) \doteq C' \log (K - N_x) + R,$$

where K is an upper limit of density.

The simulated sequences generated by these expressions, are, at least in style, very similar to real profiles and also resemble the models of population growth through time, that were proposed by Whittaker and Goodman (1979) and which were based on the integration of the logistic equation, with random changes in the parameters. This similarity is suggestive of further developments. It is natural to expect that in the center of the disturbances, where reactants are well mixed, peaks of production would appear, generated largely by r-selected organisms; and that in the peripheral areas, the environment can be qualified as more "benign" and able to support populations of K-selected species. The basic picture is always one of contrast between pointlike, discontinuous, and unpredictable events, and the ensuing ecological succession which is more gradual, predictable and ubiquitous, except where interrupted. All events are amenable to spectral expression (Abbott et al. 1982).

Disturbances are essentially unpredictable and the information obtained along a transect cannot predict the message to come. It is hopeless to pretend to maximize new information about eventual changes along the direction of the transects on the basis of precedent data.

Patterns in fixed communities (e.g., benthos) can be reduced to two dimensions and be subjected to repeated measurements. Interactions between neighbors are more consistent than those occurring in plankton and probably leads to patterns with more biological significance. A scanning of the elementary units can be done in parallel runs, as with TV. It is possible also to scan in perpendicular directions, so as to detect orientation. Each scanned unit can be defined biologically, and also in terms of color and brightness. As a first order cut at statistical analysis the densities of the various categories of elements are used to calculate their relative frequencies and the diversity. In the second iteration the degree of clumping is ascertained. Probably these measures will prove enough to suitably classify the communities by comparing the resultant values with the corresponding measures obtained from color pictures of standardized surfaces of benthos. Probably what is known about the dynamics of plankton will be of some help in trying to evaluate how remote from disturbance the observed pattern lies.

Coda

We can choose to view the ecosystem synoptically as a set of "symbols" (species, etc.)

distributed in space. Judging from certain associations among the symbols from past experience with community processes, we will probably discover some rules of proximity between symbols, and be able to visualize the historical development of the given situation. The objective is to identify and describe the processes that have led to the garbled message in the actual string of symbols.

Part of the difficulty in learning the language of nature is the confusion introduced by the continual resetting (obliteration, perturbation) of old systems and their replacement by systems with novel styles. Systems with more information in the different parts, with lower covariance in the distribution of components or reactants, and with more individuality in each component tend to grow faster and prevail. When they become large, they attain a disproportionate capacity as carriers of information at all levels. This may be an equivalent statement of the theory of natural selection. But the capacity to learn tends to block the evolution of other properties of the organism, because it diminishes their importance to survival in favor of a better capacity for learning. Fitness, as the term is ordinarily used or operationally defined, has almost nothing to do with the accumulation and utilization of information and is simply a statement of what has happened, often giving exaggerated importance to the survival strategies. The only fitness is to keep playing.

In a tentative way, we could propose a rule that should accept or reject new species or mutant in an ecosystem, according to whether or not the expansion of the new element leads to a decrease of the energy exchanged per unit of preserved biomass (or information). This would provide another rather unorthodox view of fitness, in that fitness becomes a function of the workings of the whole ecosystem. This criterion would, of course, favor *K*-selection.

Prediction is made more difficult by the fact that the most important events are usually the most improbable. All perturbations are introduced by external energy, and the whole play is between such disturbances or disasters, and the endogenous trend in all self-organizing systems to convert decaying energy into organization. This process usually slows down and ends in configurations that appear to be in steady state for rather short times. They differ locally in aspect and degree of organization from the general characteristics of the surrounding environment. Taken together, the different points of view that have been presented here provide some grounds for discussing the relations between the very general concepts of thermodynamics and the everyday problems encountered in the study of plankton. My opinion is that the variety of scales at which such pelagic ecosystems work should force the development of new ideas, that may be applicable to other domains of the biosphere.

References

ABBOTT, M. R., T. M. POWELL, AND P. J. RICHERSON. 1982. The relationship of environmental variability to the spatial patterns of phytoplankton biomass in Lake Tahoe. J. Plankton Res. 4: 927–941.

ACEVEDO, M. F. 1981. Non-equilibrium ecology: Chronic and impulsive disturbances. Manuscript 9 p. (In press)

CENTRELLA, J., AND A. L. MELOTT. 1983. Three-dimensional simulation of large-scale structure in the Universe. Nature 305: 196–198.

HANSON, F. B., AND H. C. TUCKWELL. 1981. Logistic growth with random density independent disasters. Theoret. Pop. Biol. 19: 1–18.

ISAACS, J. D. 1972. Unstructured marine food webs and "pollutant analogues". Fish. Bull. 70: 1053–1059.

LANGE, G. D., AND A. C. HURLEY. 1975. A theoretical treatment of unstructured food webs. Fish. Bull. 73: 378–381.

LEEUWEN, C. G. VAN. 1965. Gorteria. 2: 93–105. 1966. Wentia 15: 25–46.

LEVIN, S. A., AND L. A. SEGEL. 1976. Hypothesis for origin of planktonic patchiness. Nature 259: 659.

LIAPUNOV, M. A. 1892. (Problème général de la stabilité du mouvement. Ann. de Toulouse 9(2): 203–474. (translated from Russian)

McNaughton, S. J. 1978. Stability and diversity of ecological communities. Nature 274: 251–253.

Mandelbrot, B. B. 1983. The fractal geometry of nature. W. H. Freeman & Co., New York, NY. 468 p.

Margalef, R. 1968. Perspectives in ecological theory. Chicago Univ. Press, Chicago – London, 111 p.

1969. Estudios sobre la distribución a pequeña escala del fitoplancton marino. Mem. R. Acad. Cienc. Art. Barcelona 40: 1–22.

1974. Ecologia. Ediciones Omega, Barcelona, 951 pp.

1978a. Diversity, p. 251–260. *In* A. Sournia [ed.] Phytoplankton manual, Unesco, Paris.

1978b. Life-forms of phytoplankton as survival alternatives in an unstable environment. Oceanol. Acta 1: 493–510.

1979. The organization of space. Oikos 33: 152–159.

1980. La biosfera entre la temodinámica y el juego. Ediciones Omega, Barcelona. 236 p.

Margalef, R., and E. Gutiérrez. 1983. How to introduce connectance in the frame of an expression for diversity. Am. Nat. 121: 601–607.

Patten, B. C. 1982. Indirect causality in ecosystems: its significance for environmental protection. *In* W. T. Mason and S. Iker [ed.] Research on fish and wildlife habitat, U.S. Environ. Prot. Agency.

Rejmánek, M., P. Kindlmann, and L. Leps. 1983. Increase of stability with connectance in model competition communities. J. Theor. Biol. 101: 649–656.

Steele, J. H. 1974. Spatial heterogeneity and population stability. Nature, 248: 83.

Turing, A. M. 1950. Computing machinery and intelligence. Mind 59: 433–460.

Ulanowicz, R. E. 1980. An hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

Whittaker, R. H., and D. Goodman. 1979. Classifying species acording to their demographic strategy. Am. Nat. 113: 185–200.

Information and Complexity

ROBERT ROSEN

Department of Physiology and Biophysics, Dalhousie University, Halifax, N.S. B3H 4H7

Introduction

My task, as I understand it, is to discuss some of the theoretical tools that may be brought to bear on specific problems of common concern, using the concept of "information" as the point of departure.

Before turning directly to this, I feel it may be helpful to begin with a few words concerning the nature and purpose of theory itself. I do this, not only because many may differ widely about it, but mainly because my views about theory are integrally related with my approach to it. Thus, the more formal remarks to follow later may be easier to understand if prefaced by some metatheoretical prologue.

I think it can be commonly agreed that no one, be he experimenter, observer or theoretician, does science at all without believing that nature obeys laws or rules, and that these natural regularities can be at least partly grasped by the mind. That nature obeys laws is often subsumed under the notion of *causality*. The articulation of these causal laws or relationships means, in brief, that one can establish a correspondence, or dictionary, between events in the world and propositions in some appropriate language, such that the causal relations between events are exactly mirrored by implication relations between corresponding propositions. In other words, the very concept of natural law presupposes (a) that causal order exists, and (b) that it can be imaged by logical order. The theorist's job, as I see it, is essentially to bring causal order and logical or implicative order into congruence.

From this point of view, most, if not all, of the deep questions in science translate into the questions about the kinds of logical systems which can be images of causal order. For instance, the thrust of mechanics (classical or quantum) is to establish a class of mathematical systems (more broadly, dynamical systems) which can be the images of at least certain kinds of natural or material systems. On the other hand, the main thrust of Relativity, or of the Laws of Thermodynamics, is to circumscribe the class of mathematical systems which can image physical reality.

What I will argue in the following pages is that the class of mathematical images of causal order which we are used to, and which we have inherited essentially unchanged from the time of Newton, is in fact a very special class. Limiting theoretical science to the class of material systems which have images in this class thus severely circumscribes it, thereby restricting our science to what I shall call *simple systems* or *mechanisms*. I shall employ informational arguments to motivate these assertions, and also to indicate how we may effectively move into a larger world of *complex systems*, which simple ones can only approximate, locally and temporarily.

As may be imagined, we will be grappling with a large subject, and in these few pages I can only touch on a few of the most important points. But I hope that this will suffice to indicate the general trend, and to hint at the manifold practical implications in the theoretical development.

Interrogation

I shall introduce the rather wide-ranging considerations which follow with a discussion of the concept of *information* and its role in scientific discourse. Ever since Shannon began

to talk about "Information Theory" (by which he meant a probabilistic analysis of the deleterious effects of propagating signals through "channels"; cf. Shannon 1949) this concept has been relentlessly analyzed and re-analyzed. The time and effort expended on these analyses must surely rank as one of the most unprofitable investments in modern scientific history; not only has there been no profit, but the currency itself has been debased to worthlessness. Yet, in biology, for example, the terminology of information intrudes itself insistently at every level; code, signal, computation, recognition. It may be that these informational terms are simply not scientific at all; that they are an anthropomorphic stopgap; a *facon de parler* which merely reflects the immaturity of biology as a science, to be replaced at the earliest opportunity by the more rigorous terminology of force, energy, and potential which are the province of more mature sciences (i.e., physics) in which "information" is never mentioned. Or, it may be that the informational terminology which seems to force itself upon us bespeaks something fundamental; something which is missing from physics as we now understand it. I shall pursue this latter viewpoint and see where it leads us.

In human terms, information is easy to define; it is anything which is or can be the answer to a question. Therefore, I shall preface my more formal considerations with a brief discussion of the status of interrogatives, in logic and in science.

The amazing fact is that interrogation is not ever a part of formal logic, including mathematics. The symbol "?" is not a logical symbol, as for instance are "$\vee$", "$\wedge$", "$\exists$", or "$\forall$"; nor is it a mathematical symbol. It belongs entirely to informal discourse; and, as far as I know, the purely logical or formal character of interrogation has never been investigated. Thus, if "information" is indeed connected in an intimate fashion with interrogation, it is not surprising that it has not been formally characterized in any real sense. There is simply no existing basis on which to do so.

I do not intend to go deeply here into the problem of extending formal logic (always including mathematics in this domain) so as to include interrogatories. What I want to suggest here is a relation between our informal notions of interrogation and the familiar logical operation "$\Rightarrow$"; the conditional, or the implication operation. Colloquially, this operation can be rendered in the form "If A, then B". My argument will involve two steps. First, I will argue that *every* interrogative can be put into a kind of conditional form:

If A, then B?

(where B can be an indefinite pronoun like "who", "what", etc., as well as a definite proposition). Second, and most important, I will argue that every interrogative can be expressed in a more special conditional form, which can be described as follows. Suppose I know that some proposition of the form

If A, then B

is true. Suppose I now change or vary A; i.e., replace A by a new expression, which I will call δA. The result will be an interrogative, which I can express as

If δA, then δB?

Roughly, I am treating the true proposition "If A, then B" as a reference, and I am asking what happens to this proposition if I replace the reference expression A by the new expression δA. I could of course do the same thing with B in the reference proposition; replace it by a new proposition δB and ask what happens to A. I assert that every interrogative can be expressed this way, in what I shall call a *variational form*.

The importance of these notions for us will lie in their relation to the external world; most particularly in their relation to the concept of *measurement*, and to the notions of

causality to which they become connected when a formal or logical system is employed to represent what is happening in the external world (i.e., to describe some physical or biological system or situation).

Before doing this, I want to motivate the two earlier assertions regarding the expression of arbitrary interrogatives in a kind of conditional form. I will do this by considering a few typical examples, and leaving the rest to the reader for the moment.

Suppose I consider the question,

"Did it rain yesterday?"

First, I will write it in the form,

"If (yesterday), then (rain)?",

which is the first kind of conditional form described above. To find the variational form, I presume I know that some proposition like,

"If (today), then (sunny)"

is true. The general variational form of this proposition is,

"If δ(today), then δ(sunny)?"

In particular, then, if I put

δ(today) = (yesterday)

and

δ(sunny) = (rain),

I have indeed expressed my original question in the variational form. A little experimentation with interrogatives of various kinds taken from informal discourse (of great interest are questions of classification, including existence and universality) should serve to make manifest the generality of the relation between interrogation and the implicative forms described above. Of course, this cannot be *proved* in any logical sense, since, as noted above, interrogation sits outside logic.

It is clear that the notions of observation and experiment are closely related to the concept of interrogation. That is why the results of observation and experiment (i.e., data) are so generally regarded as being information. In a formal sense, simple observation can be regarded as a special case of experimentation. Intuitively, an observer simply determines what *is*, while an experimenter systematically perturbs what is, and then observes the effects of this perturbation. In the conditional form, then, an observer is asking a question which can generally be expressed as:

"If (initial conditions), then (meter reading)?"

In the variational form, this question may be formulated as follows: assuming the proposition,

"If (initial conditions = 0), then (meter readings = 0)"

is true (this establishes the reference, and corresponds to calibrating the meters), our question becomes,

"If δ(initial conditions = 0), then δ(meter readings = 0)?",

where simply,

$$\delta(\text{initial conditions} = 0) = (\text{initial conditions})$$

and

$$\delta(\text{meter readings} = 0) = (\text{meter readings}).$$

The experimentalist essentially takes the results of observation as his reference, and thus, basically asks the question which in variational form is just

"If δ(initial conditions), then δ(meter readings)?"

The theoretical scientist, on the other hand, deals with a different class of question; namely, with the questions which arise from assuming a δB (which may be B itself) and asking for the corresponding δA. This is a question which an experimentalist cannot approach directly, not even in principle. It is mainly the distinction between the two kinds of questions that marks a significant difference between experiment and theory, as well as the difference between the explanatory and predictive roles of theory itself. Clearly, if we give δA and ask for the consequent δB, we are predicting, whereas if we assume a δB and ask for the antecedent δA, we are explaining.

It should be noted that exactly the same duality arises in mathematics and logic themselves; i.e., in purely formal systems. Thus a mathematician can ask (*informally*): If (I make certain assumptions), then (what follows)? Or, he can start with a conjecture, and ask: If (Fermat's Last Theorem is true), then (what initial conditions must I assume to explicitly construct a proof)? The former is analogous to prediction, the latter to explanation.

When formal systems (i.e., logic and mathematics) are used to construct images of what is going on in the world, then interrogations and implications become associated with ideas of causality. Indeed, the whole concept of natural law depends precisely on the idea that causal processes in natural systems can be made to correspond with implication in some appropriate descriptive inferential system (e.g., Rosen 1983, where this theme is developed at great length).

But the concept of causality is itself a complicated one; this fact has been largely overlooked in modern scientific discourse, to its detriment. That causality is complicated was already pointed out by Aristotle. To Aristotle, all science was animated by a specific interrogative: why? He said explicitly that the business of science was to concern itself with "the why of things". In our language, these are just the questions of *theoretical* science: if (B), then (what A)? Whence we can say B *because* A, or, in the variational form, δB *because* δA.

But Aristotle argued that there were four distinct categories of causation; four ways of answering the question *why*. These categories he called *material cause*, *formal cause*, *efficient cause*, and *final cause*. These categories of causation are not interchangeable. If this is so (and I will argue below that indeed it is), then there are correspondingly *different kinds of information* associated with different causal categories (see also Atlan, this volume). These different kinds of information have been confused, mainly because we are in the habit of using the same mathematical language to describe all of them; and it is from these inherent confusions that much of the ambiguity and murkiness of the concept of information ultimately arises. Indeed, we can say more than this: the very fact that the same mathematical language does not (in fact, cannot) distinguish between essentially distinct categories of causation means that the mathematical language we have been using is in itself somehow fundamentally deficient, and that it must be extended by means of supplementary structures to eliminate those deficiencies.

Information

I have said above that information is, or can be, the answer to a question, and that a question can generally be put into what is called the variational form: If δA, then δB?. This premise is going to serve as the connecting bridge between "information" and the Newtonian paradigm I have described. In fact, it has played an essential role in the historical development of Newtonian mechanics and its variants, under the rubric of *virtual displacements*.

In mechanics, a virtual displacement is a small, imaginary change imposed on the *configuration* of a mechanical system, while the impressed forces are kept fixed. The animating question is: "If such a virtual displacement is made under given circumstances, then what happens?" The answer, in mechanics, is the well-known *Principle of Virtual Work*: If a mechanical system is in equilibrium, then the virtual work done by the impressed forces as a result of the virtual displacement must vanish. This is a static (equilibrium) principle, but it can readily be extended from statics to dynamics where it is known as *D'Alembert's Principle*. In the dynamic case, it leads directly to the differential equations of motion of a mechanical system when the impressed forces are known. Details can be found in any text on classical mechanics.

In what follows, I am going to explore the effect of such virtual displacements on the apparently more general class of dynamical systems of the form,

$$(1) \quad dx_i/dt = f_i(x_1, \ldots, x_n), \qquad i = 1, \ldots, n.$$

(In fact, however, there is a close relationship between the general dynamical systems (1) and those of Newtonian mechanics. Indeed the former systems can be regarded as arising out of the latter by the imposition of a sufficient number of non-holonomic constraints.)

As I have already noted, the language of dynamical systems, like that of Newtonian mechanics, does not include the word "information". Rather, the study of such systems revolves around the various concepts of *stability*. However, in one of his analyses of oscillations in chemical systems, J. Higgins (1967) drew attention to the quantities,

$$u_{ij}(x_1, \ldots, x_n) = \partial/\partial x_j(dx_i/dt).$$

These quantities, which he called "cross-couplings" if $i \neq j$ and "self-couplings" if $i = j$, arise in an essential way in the conditions which govern the existence of oscillatory solutions of (1). In fact, it turns out that it is not so much the magnitudes, as the signs, of these quantities which are important. In order to have a handy way of talking about the signs of these quantities, he proposed that we call the jth state variable x_j an *activator* of the ith, at a state $(x_1^0, \ldots, x_n^0)$, whenever the quantity,

$$u_{ij}(x_1^0, \ldots, x_n^0) = \frac{\partial}{\partial x_j}\left(\frac{dx_i}{dt}\right)_{(x_1^0, \ldots, x_n^0)} > 0,$$

and an *inhibitor* whenever

$$u_{ij}(x_1^0, \ldots, x_n^0) < 0.$$

Now activation and inhibition are *informational* terms. Thus, Higgins' terminology provides an initial hint about how dynamical language might be related to informational language, through the Rosetta Stone of stability.

Now let us see what Higgins' terminology amounts to. If x_j activates x_i at a state, it means that a (virtual) increase in x_j increases the *rate of change* of x_i, or alternatively, that a (virtual) decreases of x_j decreases the rate of change of x_i. It is eminently and intuitively reasonable that this is what an activator should do. Conversely, if x_j inhibits x_i at a state, it means that an increase in x_j decreases the rate of change of x_i, etc.

Thus the n^2 functions $u_{ij}(x_1, \ldots, x_n)$, $i, j = 1, \ldots, n$ constitute a kind of informational description of the dynamical system (1), which I have elsewhere (Rosen 1979) called an *activation–inhibition pattern*. As I have noted, such a pattern concisely represents the answers to the vartiational questions: "If we make a virtual change in x_j, what happens to the rate of production of x_i?".

There is no reason to stop with the quantities u_{ij}. We can, for instance, go one step further, and consider the quantities

$$u_{ijk}(x_1, \ldots, x_n) = \partial/\partial x_k(\partial/\partial x_j(dx_i/dt)).$$

Intuitivcly, these quantities measure the effect of a (virtual) change in x_k on the *extent* to which x_j activates or inhibits x_i. If such a quantity is positive at a state, it is reasonable to call x_k an *agonist* of x_j with respect to x_i; if it is negative, an *antagonist*. That is, if u_{ijk} is positive, a (virtual) increase in x_k will increase or facilitate the activation of x_i by x_j, etc. The quantities u_{ijk} thus define another layer of informational interaction, which we may call an *agonist–antagonist pattern*.

We can iterate this process, in fact to infinity, producing at each state r a family of n^r functions $u_{ij\ldots r}(x_1, \ldots, x_n)$. Each layer in this increasing sequence describes how a (virtual) change of a variable at that level modulates the properties of the preceding level.

In the above considerations, I have treated only the effects of virtual changes in state variables x_j on the velocities dx_i/dt at various informational levels. We could similarly consider the effects of virtual displacements at these various levels on the second derivatives d^2x_i/dt^2 (i.e., on the *accelerations* of the x_i), on the third derivatives d^3x_i/dt^3, and so on. Thus, we have a doubly infinite web of informational interactions, defined by functions,

$$u^m_{ijk\ldots r}(x_1, \ldots, x_n) = \frac{\partial}{\partial x_r}\left(\ldots \frac{\partial}{\partial_{x_j}}\left(\frac{d^m x_i}{dt^m}\right)\ldots\right).$$

If we start from the dynamical equations (1), then nothing new is learned from these circumlocutions, beyond perhaps a deeper insight into the relations between dynamical and informational ideas. Indeed, given any layer of informational structure, we can pass to succeeding layers by mere differentiation, and to antecedent layers by mere integration. Thus, knowledge of any layer in this infinite array of layers determines all of them, and, in particular, the dynamical equations themselves. For instance, if we know the activation–inhibition pattern $u_{ij}(x_1, \ldots, x_n)$, we can reconstruct the dynamical equations (1) through the relations,

$$(2) \quad df_i = \sum_{j=1}^{n} u_{ij}dx_j$$

(note in particular that the differential form on the right-hand side is like a generalized *work*), and then equating the function, $f_i(x_1, \ldots, x_n)$, so determined to the rate of change dx_i/dt of the ith state variable.

However, *our ability to do all this depends in an absolutely essential way on the exactness of the differential forms which arise at every level of our web of informational interaction*, and which relate each level to its neighbors. For instance, if the forms in (2) are not exact, there are no functions $f_i(x_1, \ldots, x_n)$ whose differentials are given by (2), and hence *no rate equations of the form (1)*. In fact, in such a situation, the simple relations between the levels in our web (namely, that each level is the derivative of the preceding level and the integral of the succeeding one) breaks down completely; the levels may become independent of each other, and must be posited separately. Thus, for instance, two systems could have the same activation–inhibition patterns, but vastly different agonist–antagonist patterns, and hence manifest entirely different behaviors.

Just to fix ideas, let us see what is implied by the requirement that the differential forms,

$$\sum_{j=1}^{n} u_{ij} dx_j,$$

defined by the activation–inhibition pattern be exact. The familiar necessary conditions for exactness take the form,

$$\frac{\partial}{\partial x_k} u_{ij} = \frac{\partial}{\partial x_j} u_{ik},$$

for all $i, j, k = 1, \ldots, n$. Intuitively, these conditions mean precisely that *the relations of agonism and activation are entirely symmetrical* (commutative); that x_k as an agonist of the activator x_j is exactly the same as x_j as an agonist of the activator x_k. And likewise for all other levels.

Clearly, in informational terms such situations are extremely degenerate because the requirement of exactness is highly nongeneric for differential forms. Thus, these very simple considerations suggest a most radical conclusion: *that the Newtonian paradigm, with its emphasis on dynamical laws, restricts us from the outset to an extremely special class of systems, and that the most elementary informational considerations force us out of that class*. I shall explore some of the implications of this situation in the subsequent section.

Meanwhile, let us consider some of the ramifications of these informational ideas, which hold even within the confines of the Newtonian paradigm. These will concern the distinctions between environment, phenotype, and genome, the relations of these distinctions to different categories of causation, and the correspondingly different types of information which these causal categories determine.

First, I have asserted elsewhere (Rosen 1983), that according to the Newtonian paradigm, every relation between physical magnitudes (i.e., every equation of state) can be represented as a genome-parameterized family of mappings,

$$f_g : E \to P,$$

from environments to phenotypes. It is worth noting that, in particular, every dynamical law or equation of motion is of this form, as can be seen by writing,

$$(3) \quad d\vec{x}/dt = f_g(\vec{x}, \vec{a}).$$

Here, in traditional language, $\vec{x}$ is a vector of states, $\vec{a}$ is a vector of "external controls" (which together with states constitutes *environment*), and the phenotype is precisely the tangent vector $d\vec{x}/dt$ attached to the state $\vec{x}$. In this case, then, tangent vector (or phenotype) constitutes *effect*; genome g is identified with formal cause, state x with material cause, and the operator $f_g(\ldots, \vec{a})$ with efficient cause.

By analogy with the activation–inhibition networks and their associated informational structures which were described above, I am going to consider formal quantities of the form,

$$(4) \quad \frac{\partial}{\partial(\text{cause})} \left(\frac{d}{dt} (\text{effect}) \right).$$

As always, such a formal quantity represents an answer to a question: If (cause is varied), then (what happens to effect)? This is exactly the same question I asked in connection with the definition of activation–inhibition networks and their correlates, but now set into the wider context to which my analysis of the Newtonian paradigm has led us. That is, we may now virtually displace *any* magnitude which affects our relation (3), whether it be a

genomic magnitude, an environmental magnitude, or a state variable. In a precise sense, the effect of such a virtual displacement is measured by the quantity (4).

In particular, then, it follows that there are indeed different *kinds* of information. What kind of information we are dealing with depends on whether we apply our virtual displacement to a genomic magnitude (associated with formal cause), an environmental magnitude (efficient cause), or a state variable (material cause). Initially, then, we can formally distinguish at least the following three cases:

1. Genomic information:

$$(5)\quad \frac{\partial}{\partial(\text{genome})}\left(\frac{d}{dt}\,(\text{effect})\right);$$

2. Phenotype information:

$$\frac{\partial}{\partial(\text{state})}\left(\frac{d}{dt}\,(\text{effect})\right);$$

3. Environmental information:

$$\frac{\partial}{(\text{control})}\left(\frac{d}{dt}\,(\text{effect})\right).$$

I shall confine myself to these three for present purposes, which generalize only the activation–inhibition patterns described above.

We now come to an important fact; namely, *the three categories defined above are not equivalent*. Before justifying this assertion, I must spend a moment discussing what is meant by "equivalent". In general, the mathematical assessment of the effects of perturbations (i.e., of real or virtual displacements) is the province of *stability*. The effect on subsequent dynamical behavior of modifying or perturbing a system state is the province of Lyapunov stability of dynamical systems; the effect of perturbing a control is part of control theory; the effect of perturbing a genome is the province of structural stability. To fix ideas, let us consider genomic perturbations, or *mutations*. A virtual displacement applied to a genome g replaces the initial mapping determined by g, f_g, with a new mapping $f_{g'}$. Mathematically, we say that the two mappings f_g, $f_{g'}$ are equivalent (or similar, or conjugate), if there exist appropriate transformations,

$$\alpha : E \to E,$$

and

$$\beta : P \to P,$$

such that the diagram,

$$\begin{array}{ccc} E & \xrightarrow{f_g} & P \\ \alpha\downarrow & & \downarrow\beta \\ E & \xrightarrow[f_{g'}]{} & P \end{array}$$

commutes; i.e., if

$$\beta(f_g(e)) = f_{g'}(\alpha(e)),$$

for every e in E. Intuitively, this means that a mutation $g \to g'$ can be offset, or annihilated, by imposing suitable *coordinate transformations* on the environments and phenotypes. Stated yet another way, a virtual displacement of a genome can always be

counteracted by corresponding displacements of environment and phenotype so that the resultant variation on effect vanishes.

I have elsewhere (Rosen 1978) shown at great length that this commutativity may not always obtain; i.e., that there may exist genomes which are bifurcation points. In any neighbourhood of a bifurcating genome g, there exist genomes g' for which f_g and $f_{g'}$ fail to be conjugate.

With this background, I return to the question of whether the three kinds of information (genomic, phenotypic, and environmental) which I have defined above are equivalent. Intuitively, equivalence would mean that the effect of a virtual displacement, δg, of genome, say with everything else held fixed, could equally well be produced by a virtual displacement, δa, of environment, or by a virtual displacement, δp, of phenotype. Stated another way, the effect of virtual displacement, δg, of genome can be annihilated by virtual displacements, $-\delta a$ and $-\delta p$, of environment and phenotype, respectively. This is simply a restatement of the definition of conjugacy or similarity of mappings.

If all forms of information were equivalent, it would follow that there could be no bifurcating genomes. I note in passing that the assumption of equivalence of the three kinds of information defined above creates terrible ambiguities when it comes to the *explanation* of particular effects. I will not consider that aspect here, except to note that it is perhaps very fortunate for us that, as we have seen, they are not equivalent.

Let us look at one immediate consequence of the non-equivalence of genomic, environmental, and phenotypic information, and of the considerations which culminate in that conclusion. Long ago (cf. von Neumann 1951; Burks 1966) von Neumann proposed an influential model for a "self-reproducing automaton", and subsequently, for automata which "grow" and "develop". This model was based on a famous theorem of Turing (1936) establishing the existence of a universal computer (universal Turing machine). From the existence of such a universal computer, von Neumann argued that there must also exist a universal constructor. Basically, he argued that computation (i.e., following a program) and construction (following a blueprint) are both algorithmic processes, and that anything holding for one class of algorithmic processes necessarily holds for any other class. This universal constructor formed the central ingredient of the "self-reproducing automaton".

Now a computer acts, in the language developed above, through the manipulation of efficient cause. A constructor, if the term is to have any shred of its intuitive meaning, must essentially manipulate material cause. The inequivalence of the two categories of causality, in particular as it manifested by the non-equivalence of environmental and phenotypic information, means that we cannot blithely extrapolate from results pertaining to efficient causation into the realm of material causation. Indeed, in addition to invalidating von Neumann's specific argument, we learn that great care must be exercised in general when arguing from purely logical models (i.e., from models pertaining to efficient cause) to any kind of physical realization, such as developmental or evolutionary biology (which, as noted, pertain to material cause).

Thus, we see how significant are the impacts of informational ideas, even within the confines of the Newtonian paradigm. In this paradigm, as I have shown, the categories of causation are essentially segregated into separate packages. I will now turn to the question of what happens when we leave the comforting confines of that paradigm.

Some Consequences

I am going to call any natural system for which the Newtonian paradigm is completely valid a *simple system*, or *mechanism*. In this terminology, then, a *complex system* is one

which falls outside this paradigm for any reason. We have already seen a hint of such systems above; e.g., in systems whose activation–inhibition patterns u_{ij} do not give rise to exact differentials.

This view of complexity as something transcending our familiar modes of system description is quite different from the more conventional view, which identifies complexity with what is merely technically difficult within those modes. In this conventional view, complexity is just one more system property, measured by a number (e.g., the dimensionality of a state space, the length of an algorithm, the time or energy involved in some computation, etc.). Although I cannot take the time to discuss the point here, my approach to complexity is more in accord with the heuristic one, that complex systems do unexpected or counter-intuitive things.

Now, let us see where information fits in. We recall once more that "information" is an actual or potential response to an interrogative, and that every interrogative can be put into a variational form. The Newtonian paradigm asserts, among other things, that the answers to such interrogatives follow from dynamical law superimposed on manifolds of states. In their turn, such dynamical laws are special cases of what I have elsewhere called *equations of state*, which serve to link or relate the values of system observables. Indeed, the observable is the basic ingredient in our approach to system representation (cf. Rosen 1978); it is the connecting link between the world of natural phenomena and the entirely different world of formal systems, which we use to describe and explain.

However, the kinds of considerations I have developed above suggests that this world is not enough. We require also a world of variations, increments, and differentials of observables. It is true that every linkage between observables implies a corresponding linkage between differentials, but as we have seen, the converse is not true. We are thus led to the notion that a differential relation is a generalized linkage, and that a differential form is a kind of generalized observable. A differential form which is not the differential of an observable is thus an entity which assumes no definite numerical value (as an observable does), but which can be incremented.

If we do think of differential forms as generalized observables, then we must correspondingly generalize the notion of equation of state. A generalized equation of state thus becomes a linkage or relation between ordinary observables and differentials or generalized observables. Such generalized equations of state are the vehicles which answer questions of the variational form: If δA, then δB?

But as I have repeatedly noted, such generalized equations of state do not generally follow from systems of dynamical equations, as they do in the Newtonian paradigm. Thus, we must find some alternative way of characterizing a system of this kind. Here is where the informational language which I have introduced above comes to the fore. Let us recall, for instance, how I defined the activation–inhibition network. I described a family of functions u_{ij} (i.e., of observables) which could be thought of in the dynamical context as modulating the effect of an increment, dx_j, on that of another increment, df_i. That is, the values of each observable, u_{ij}, measure precisely the extent of activation or inhibition which x_j exerts on the rate at which x_i is changing.

In this language, a system falling outside the Newtonian paradigm (i.e. a complex system) can have an activation–inhibition pattern, just as a dynamical (i.e., simple) system does. Such patterns are still families of functions (observables), u_{ij}, and the pattern itself is manifested by the differential forms,

$$\omega_i = u_{ij}dx_j.$$

But in this case, there is no global velocity observable, f_i, which can be interpreted as the rate of change of x_i; there is only a velocity *increment*. It should be noted explicitly that

the u_{ij}, which define the activation–inhibition pattern, need not be functions of the x_i alone, or even depend on them at all. Thus, the differential forms which arise in this context are different from those with which mathematicians generally deal.

The next level of information is the agonist–antagonist pattern, u_{ijk}. In the category of dynamical systems this is completely determined by the activation–inhibition pattern, and can be obtained from the latter by differentiation:

$$u_{ijk} = \frac{\partial}{\partial x_k} u_{ij}.$$

In the complex world of generalized observables and linkages, the u_{ijk} may be independent of the u_{ij}, and must be posited separately. In other words, complex (non-Newtonian) systems can have identical activation–inhibition patterns but quite different agonist–antagonist patterns.

Exactly the same considerations can now be applied to every subsequent layer of the informational hierarchy; each of them is now potentially independent of the others, and must be posited separately. Hence, a complex system generally requires an *infinite* mathematical object for its description.

I cannot go into the mathematical details of the considerations sketched so briefly above. Suffice it to say that a complex system, defined by a hierarchy of informational levels of the type I have described, is quite a different kind of object than is a dynamical system. For one thing, it is quite clear that there is not such thing as a set of *states*, assignable to such a system once and for all. From this alone, we might expect that the nature of causality in such systems is vastly different than it is in the Newtonian paradigm; we shall come to this in a moment.

The totality of mathematical structures of the type defined above forms a category. In this category the class of general dynamical systems constitutes a very small subcategory. I am suggesting that the former provides a suitable framework for the mathematical imaging of complex systems, while the latter, by definition, can image only simple systems or mechanisms. If these considerations are valid (and I believe they are), then the entire epistemology of our approach to natural systems is radically altered, and it is the basic notions of information which provide the natural ingredients for the new outlook.

There is, however, a profound relationship between the category of general dynamical (i.e., Newtonian) systems, and the larger category in which it is embedded. This can only be indicated here, but it is important indeed. Namely, there is a precise sense in which an informational hierarchy can be *approximated*, locally and temporarily, by a general dynamical system. With this notion of approximation there is an associated notion of *limit*, and hence of topology. Using these ideas, it can in fact be shown that what we can call the category of complex systems is the completion, or limiting set, of the category of simple (i.e., dynamical) systems.

The fact that complex systems can be approximated (albeit locally and temporarily) by simple ones is a crucial one. It explains precisely why the Newtonian paradigm has been so successful, and why, to this day, it represents the only effective procedure for dealing with system behavior. But in general, we can also see that it can supply *only* approximations, and in the universe of complex systems this amounts to replacing a *complex* system with a *simple subsystem*. Some of the profound consequences of doing this are considered in detail in Rosen (1978).

This relationship between complex systems and simple ones is, by its very nature, without a reductionistic counterpart. Indeed, what we presently understand as "physics" is seen in this light as *the science of simple systems*. The relationship of physics to biology is thus not at all the relation of general to particular; in fact, quite the contrary. It is not

biology, but physics, which is too special. We can see from this perspective that biology and physics (i.e., contemporary physics) grow as two divergent branches from a *theory of complex systems* which as yet can be glimpsed only very imperfectly.

The category of simple systems is, however, still the only thing we know how to work with. But to study complex systems by approximating them as simple systems puts us in the position of early cartographers, who were attempting to map a sphere while armed only with pieces of planes. Locally, and temporarily, they could do very well, but globally, the effects of the topology of the sphere become progressively important. So it is with complexity; over short times and only a few informational levels, we can always make do with a simple (i.e., dynamical) picture. Otherwise, we cannot; we must continually replace our approximating dynamics by others as the old ones fail. Hence another characteristic feature of complex sytems; they appear to possess a multitude of partial dynamical descriptions, which cannot be combined into one single complete description. Indeed, in earlier work (Rosen 1978) I took this as the defining feature of complexity.

I shall add one brief word about the status of causality in complex systems, and about the practical problem of determining the functions which specify their informational levels. As we have already noted, complex systems do not possess anything like a state set which is fixed once and for all. And in fact, in complex systems, the categories of causality become intertwined in a way which is not possible within the Newtonian paradigm. Intuitively, this follows from the independence of the infinite array of informational layers which constitutes the mathematical image of a complex system. The variation of any particular magnitude connected with such a system will typically manifest itself independently in many of these layers, and thus reflect itself partly as material cause, partly as efficient cause, and even partly as formal cause in the resultant variation of other magnitudes. I feel that it is, at least in large part, this involvement of magnitudes simultaneously in each of the causal categories which makes biological systems so refractory to the Newtonian paradigm.

Also, this intertwining of the categories of causation in complex systems makes the interpretation of experimental results of the form "If δA, then δB" extremely difficult to interpret directly. If I am correct in what I have said so far, such an observational result is far too coarse as it stands to have any clear-cut meaning. To be meaningful, an experimental proposition of this form must isolate the effect of a variation δA on a single informational level, keeping the others clamped. As might be appreciated from what has been said so far, this will in general not be an easy thing to do. In other words, the experimental study of complex systems cannot be pursued with the same tools and ideas as are appropriate for simple systems.

One final conceptual remark is also in order. As I pointed out above, the Newtonian paradigm has no room for the category of final causation. This category is closely tied up with the notion of anticipation, and in its turn, with the ability of systems to possess internal predictive models of themselves and their environments, which can be utilized for the control of present actions. I have argued at great length elsewhere (cf. Rosen 1983) that anticipatory control is indeed a distinguishing feature of the organic world, and have developed some of the unique features of such anticipatory systems. In the present discussion, I have in effect shown that, in order for a system to be anticipatory, it must be complex. Thus, the entire treatment of anticipatory systems becomes a corollary of complexity. In other words, complex systems can admit the category of final causation in a perfectly rigorous, scientifically acceptable way. Perhaps this alone is sufficient recompense for abandoning the comforting confines of the Newtonian paradigm, which has served us so well over the centuries. It will continue to serve us well, provided that we recognize its restrictions and limitations as well as its strengths.

References

BURKS, A. 1966. Theory of self-reproducing automata. University of Illinois Press, Urbana, IL.

HIGGINS, J. 1967. The theory of oscillating reactions. J. Ind. Eng. Chem. 59: 18–62.

ROSEN, R. 1978. Principles of measurement and representation of natural systems. Elsevier, New York, NY.

1979. Some comments on activation and inhibition. Bull. Math. Biol. 41: 427–445.

1983. The role of similarity principles in data extrapolation. Am. J. Physiol. 244: R591–599.

SHANNON, C. 1949. The mathematical theory of communication. University of Illinois Press, Urbana, IL.

TURING, A. 1936. On computable numbers. Proc. London Math. Soc. Ser. 2, 42: 230–265.

VON NEUMANN, J. 1951. The general and logical theory of automata, p. 1–41. *In* L. A. Jeffries [ed.] Cerebral mechanisms in behavior. John Wiley, New York, NY.

VI. WORKING GROUP REPORTS

I. Hypothesis Testing and Sampling Design in Exploited Ecosystems

W. C. Leggett* (Chairman), W. R. Boynton, V. F. Gallucci, L. Johnson, R. Margalef, J. McGlade, D. Mikulecky, R. Rosen, R. O. Ulloa, and J. Watson

Mandate

Working Group I was charged with evaluating the utility of various initiatives in theoretical ecology to the development and testing of hypotheses relating to the dynamics of exploited marine ecosystems. This mandate recognizes the fact that exploited ecosystems have certain unique properties and problems, and that current methods of study and analysis are frequently inadequate to yield either reliable descriptions of the essential properties of exploited ecosystems, or accurate predictions of their future states following perturbation.

Scope

For purposes of discussion Working Group I considered the following classes of ecosystems to be within its purview:

a) Harvested (e.g., fisheries).
b) Polluted (e.g., nutrients and/or toxic substances).
c) Manipulated (e.g., mariculture).

The Working Group considered the major features distinguishing exploited from unexploited ecosystems to be the nature, frequency, and duration of the perturbations to which they are exposed. Both exploited and unexploited ecosystems are subject to regular (diel, seasonal) and stochastic (meteorologically and hydrologically driven) perturbations. Exploited ecosystems experience an additional level of man-induced perturbations. These are generally characterized by being more predictable in space and time, more directed at specific ecosystem elements (example, top predators) and more deterministic in their magnitude and scope. In addition, because of the proximity of most important production zones to major land masses, and the tendency for large-scale transfer of the effects of man's activities to the ocean environment to be via rivers and estuaries (pollution, discharge regulation, etc.), exploited ecosystems as defined tend to be coastal rather than oceanic in character.

Hypothesis Testing and Sampling Design

It is the opinion of the Working Group that exploited ecosystems offer unique opportunities for ecosystem level hypothesis testing in the classical sense (i.e., through the observation of changes in system properties in response to perturbation). The advantage of exploited ecosystems in this regard is that the timing and magnitude of these perturbations is (generally) predictable. Such perturbations can thus be properly viewed as ecosystem scale experiments. The Working Group believes that too little use has been made of this property of exploited marine ecosystems. Exploited ecosystems offer the additional advantage that long data records are frequently available for selected elements of the system.

*Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, Que. H3A 1B1.

Three generalizations, highlighted by this conference, are warranted:

1) The number of well-documented, process-oriented analyses of marine/estuarine ecosystems is severely limited. This was made evident by the repeated reference, by speakers from diverse backgrounds, to a limited number of specific ecosystems (North Sea, Crystal River, etc.).

2) Those process-oriented descriptions which do exist are highly simplified. Even the most complex (Crystal River) deals primarily with the nekton, organisms at all other trophic levels being highly aggregated in the analysis.

3) The non-linear functional components of these ecosystems are only marginally understood.

These restrictions manifest themselves in the currently available analyses of the responses of exploited ecosystems to perturbation. Studies of ecosystem level responses to nutrient additions have, for example, focused primarily on the plankton, while observations on the responses to exploitation of fisheries have concentrated almost exclusively on fishes. In comparison to the fish and plankton, very little effort has been directed towards the benthos. This is judged to be an important gap in our knowledge.

There is also a sense in the Working Group that the new insights that might be gained by the application of ecosystem theory and new analytical techniques to marine ecosystems are currently hampered by a lack of diversity and depth in the available process-oriented descriptions of these ecosystems. The committee wishes to stress, however, the distinction between the shortage of these ecosystem descriptions, and the shortage of data. There is a strong consensus that extensive bodies of data exist, and that before new field exercises are mounted a concerted effort should be made to identify, render available, and evaluate existing data which are not now a part of the published record. New data gathering should then be restricted to the closing of gaps in existing records, with particular emphasis on processes.

New Approaches

We turn now to a specific consideration of the utility of various new approaches described at this conference to hypothesis testing and sampling design in exploited ecosystems.

Information Theory, Allometry, and Particle Size

As a general rule, data unique to particular systems is less valuable or important than data which is meaningful to a class of systems (e.g., specific to the St. Lawrence estuary versus general to all estuaries). The transportability of data from one system to another is variously called the theory of similarity or scaling. Species-based data tend to be site or system specific, and hence similarity is low. Allometric and particle size descriptions, in contrast, have been shown to be effective in identifying certain similarity classes (Peters 1983). Such data are thus information-rich.

The Working Committee considers the identification of similarity classes through the use of allometric, particle size, and like techniques to be useful for two reasons:

1) Similarity generates an equivalence relation in a class of systems being compared. Any two systems in a single similarity class are intertransformable in a canonical way, and hence data about any system in that similarity class can be transformed into corresponding data about any other system in the same class. This is a powerful consideration when applied to our earlier observations that we may be data rich but description poor. If

similarity classes can be effectively identified, and data safely transferred from one system to another, the potential for greatly increasing the ecosystem level data available, without extensive additional field sampling, is enormous.

2) Perturbation of a system S_1 (exploitation) results in its conversion to a different system S_2. If S_2 is in the same similarity class as S_1 it can be rescaled (as in example 1) to annihilate the effect. If, however, S_2 represents a different similarity class, no rescaling will offset the perturbation and new controls must be superimposed. The potential for the testing of hypotheses (predictions) based on extrapolation witnin similarity classes is great. Concurrent developments of measures of similarity and approaches to scaling are, therefore, strongly recommended.

A central question in the development of a predictive theory of the dynamics of exploited ecosystems, and in the establishment of sensitivity classes, is the extent to which perturbations are propagated in ecosystems. Platt (1985) has suggested two scenarios: (1) that perturbations will propagate through the entire size spectrum with few modifications and (2) that perturbations will be damped during progression. Both possibilities were presented as responses to perturbations generated in the lower size classes. It is equally important (and feasible) to assess the extent to which perturbations in the larger size classes can be propagated backwards toward the lower size classes (early work by Dodson and Brooks (1965) suggest that it may). Exploited ecosystems offer ideal environments in which to test these propositions. The Working Committee favours such tests.

A second question of relevance to the evaluation of similarity states in exploited ecosystems is the nature of system responses to perturbation. It has, for example, generally been observed that stressed ecosystems respond by accelerating turnover. It is proposed that this generalization be further evaluated and, in addition that the relationship between particle size distribution, allometric states, and observed changes in turnover rates be determined.

Thermodynamics

The Working Group concludes that the utility of direct applications of thermodynamic principles to hypothesis testing in exploited marine ecosystems is currently questionable. This is because the dynamic nature of the marine ecosystem virtually ensures departure from the equilibrium state assumption of the theory. The Working Group does not, however, preclude the possibility of significant advances in this area in the future. The applicability of the theory should, however, be more fully evaluated in smaller, stable systems such as arctic lakes before application to large-scale marine systems is attempted.

Flow Analysis and Network Thermodynamics

It is the consensus of the Working Committee that both flow analysis and network thermodynamics hold considerable potential for the study of exploited ecosystems. In particular, the committee was impressed with their possible use in sampling design. The utility of these and similar procedures is considered to be primarily as guides to the identification of those variables most sensitive to perturbation, and most influential to ecosystem output. The Committee also holds that such approaches could prove useful in assessing the validity of assumptions (hypotheses) concerning system structure and flow dynamics between units, and the identification of different similarity classes. Continued development and application of such approaches to marine ecological data and sampling design is encouraged. We would caution, however, that these observations should not be

taken as support for a repeat of the International Biological Program experience where large-scale models were developed on the basis of minimal knowledge of ecological processes.

Statistical Mechanics

The Working Committee concurs with Dr. Trainor's (1985) refreshingly honest evaluation of the potential utility of Statistical Mechanics in the evaluation of marine ecosystems. The assumptions of Statistical Mechanics remain inconsistent with the realities of marine ecosystems. Moreover, the higher-level laws essential to its application have not yet been fully developed in Ecology. Significant developments are likely to be forthcoming in both fields, but until they occur, the methodology of Statistical Mechanics would appear to have little to contribute to hypothesis testing in marine ecosystems.

Space and Time

While this subject was not specifically addressed at this conference, the Working Group wishes to highlight the importance it places on the identification of appropriate spatial and temporal scales in the design of studies. This question has been addressed in significant detail in Legendre and Demers (1984) and will not be amplified here. It is sufficient to state that such considerations are of the utmost importance in dynamic systems of which marine ecosystems are outstanding examples. Network thermodynamics appears to have considerable potential as a method of a priori identifying appropriate temporal and spatial scales of observation, provided some information on the system under investigation is either directly or indirectly available.

References

Brooks, J. L., and S. J. Dodson. 1965. Predation, body size, and composition of plankton. Science 150: 28–35.

Legendre, L., and S. Demers. 1984. Towards dynamic biological oceanography and limnology. Can. J. Fish. Aquat. Sci. 41: 2–19.

Peters, R. H. 1983. The ecological implications of body size. Cambridge Univ. Press. 329 p.

Platt, T. 1985. Structure of the marine ecosystem: its allometric basis. *In* R. E. Ulanowicz and T. Platt [ed.] Ecosystem theory for biological oceanography. Can. Bull. Fish. Aquat. Sci. 213. (This volume)

Trainor, L. E. H. 1985. Statistical mechanics in biology — applications to ecology. *In* R. E. Ulanowicz and T. Platt [ed.] Ecosystem theory for biological oceanography. Can. Bull. Fish. Aquat. Sci. 213. (This volume)

II. Ecosystem Theory in Relation to Unexploited Marine Ecosystems

J. G. FIELD* (CHAIRMAN), F. V. WULFF (RAPPORTEUR), P. M. ALLEN, M. J. R. FASHAM, J. FLOS, S. FRONTIER, J. J. KAY, W. SILVERT, AND L. TRAINOR

Introduction

Biological systems are hierarchical, having different levels of organization ranging from ecosystem, community, population, individual, ... to cell, organelle, and molecule (Webster 1979). Many processes and properties in biological systems are linked, the links tending to be stronger within one level than between the levels. One cannot infer system properties by studying only the separate components at a lower level of organization (Mann 1982). It is therefore desirable to consider properties at the macro-ecosystem level in an attempt to develop new theory and understanding at this level. There are links from the higher levels in an ecosystem to the lower levels, such as the population level, and these levels are often better understood. Nevertheless, they will not explicitly be emphasized here for lack of space.

Statistical Mechanics and Thermodynamics

The question arises as to what extent thermodynamics and statistical mechanics, which have played such a major role in achieving an understanding of physical phenomena during the past century, might provide useful approaches to analysing and understanding ecosystems. Thermodynamics, as the name implies, is classically defined as the study of the interplay between thermal effects, largely characterized by the concept of temperature, and dynamic effects such as pressure. Statistical mechanics in the conventional sense is the science of bridging the hierarchical levels of the microscopic world with macroscopic thermodynamic phenomena.

In both thermodynamics and statistical mechanics the concept of equilibrium is central, i.e., equilibrium with respect to both thermal and dynamic effects. As one moves away from equilibrium, gradients develop (the so-called generalized forces), which give rise to flows, e.g., temperature gradients give rise to heat flow, and pressure gradients in a fluid to momentum flow. Near equilibrium the forces and flows are linearly related and steady state phenomena are well-defined (Trainor, this volume). If a thermodynamic system is pushed sufficiently far from equilibrium, instabilities associated with dissipative mechanisms develop, and the system undergoes behavioral changes to new steady states, e.g., from homogeneous states to ordered states (Allen, this volume). These shifts appear to be intriguingly analogous to the rapid changes of state which have been observed (but not predicted) in marine ecosystems [e.g., Nova Scotian kelp beds (Mann 1982); anchovy/sardine flip-flops in most upwelling areas (Daan 1980)]. The study of these instabilities and of the development of so-called dissipative structures (such as convection cells) is a subject of intense investigation throughout the physical, chemical and biological sciences. The statistical mechanical analysis of such phenomena is still very much in its infancy, but much progress has been made in associated areas, such as bifurcation theory and the development of chaos in nonlinear dynamical systems (Allen, this volume). Such quasidiscontinuous changes of state in ecological systems are poorly understood and need identification.

*Department of Zoology, University of Cape Town, Rondebosch 7700, South Africa.

Biological organisms, by their very nature, are systems operating far from equilibrium, since they are maintained fundamentally by dissipative mechanisms. Therefore, conventional statistical mechanics cannot sensibly be applied to a biological system. In ecology, however, one is often dealing with large numbers of organisms, and the question arises whether thermodynamics or statistical mechanics (which involves averaging the behavior of many particles) might be applied. It is certainly useful to talk about the mass and energy flows in such systems, and even in some cases to associate "forces" (linear or otherwise) as causing these flows. But the existence of forces and flows does not guarantee the existence of properties normally associated with thermodynamic equilibrium or with near-equilibrium behavior. At present, neither thermodynamics nor statistical mechanics is sufficiently developed regarding conditions far from equilibrium as to make them immediately applicable to ecosystems. This is not to deny the usefulness of some of the tools of thermodynamics or statistical mechanics, such as systems of nonlinear differential equations and Monte Carlo simulation models, which deserve further exploration.

Information Theory

The analogies between interactions that occur between the different components in an ecosystem and those described by information theory are intriguing. Although formal descriptions of and hypotheses on the relationships between chaos, order, and information in simple systems are well-developed, their applicability to ecological systems is largely hypothetical. More specifically, we know very little about how to measure the information of ecological systems. We know that taxonomic information, although useful at the population level, is a poor descriptor of ecosystem complexity and redundancy. Measures of connectivity (i.e., in the form of energy flows and the number of connecting nodes) seem to provide a more promising way to describe the information content of ecosystems (Margalef and Gutierrez 1983), and Ulanowicz (1980) has described the organization of a flow system in terms of ascendency. The concept of adaptability, also formulated in terms of information theory, provides an impressive and all-embracing framework for viewing the evolution of organisms in relation to ecosystems and hierarchical levels of organization (Conrad 1983 and this volume). The organization of information in a community may be partially described by species-abundance curves, which can be interpreted in terms of fractal theory (Frontier, this conference; Mandelbrot 1980).

Flow Analysis

Compartmental models of ecosystems may differ in many respects, but they all incorporate fluxes of mass or energy, and the description of a system requires an understanding of these fluxes. A number of different techniques were discussed at this symposium; loop analysis (Lane), input–output analysis (Patten, Hannon), flow analysis (Fasham), information networks (Atlan, Ulanowicz), and dynamic network analysis (Mikulecky). These emerged as a family of techniques with much in common, ranging from the purely qualitative loop analysis to the simple, linear, steady-state input–output and flow analyses, which are robust, need few data and require no assumptions about the functional relationships between components. The most complex technique is dynamic network analysis, which may involve nonlinear functions describing changes in state. Different versions of the same method may be used to deduce network structure from the flows, or flows from the structure. This flow-structure duality leads to the exciting possibility of mixed ecosystem models in which data, which often are incomplete with respect to structural information and measurements of flows, may be combined to deduce missing

components. This approach also has potential for use in predictive modelling. Sensitivity analysis enables one to test the structure and flow description of a model and to prescribe the precision required of field or experimental data.

These techniques are likely to provide common descriptors of ecosystem flow/structure and of how systems change in relation to external perturbations. The lack of an encompassing ecosystem theory is partly due to the paucity of compartmental descriptions on which to base hypotheses.

Role of Modelling

The group recognized a need for conceptual and simulation modelling activities to be carried out simultaneously with experimental and field work, especially in the execution of whole-ecosystem research programmes. The initiation of model building should be started *before* experimental work or sampling is begun, so that the researcher can use the modelling process to coordinate and plan field work as well as to analyse existing data sets. Modelling is not a goal in itself, but rather an iterative process, repeated over and over again in the process of increasing one's understanding of the system being studied.

Ataxonomic Aggregations and Allometry

The group believes that one important preliminary to describing ecosystem properties is to aggregate the system components on the basis of comparable time scales (i.e., turnover times) rather than into the conventional taxonomic units, which tend to be more appropriate for the population level. Turnover times are difficult to measure, however, hence the usefulness of the allometric relation:

$$T = w^a.$$

This allometric relation allows one to estimate turnover times from a knowledge of particle sizes (Calder, this volume; Peters 1983). Particle-size distributions (p.s.d.'s) have been measured only over limited size ranges in the water column, and more rigorous measurements and analyses of p.s.d.'s are required, particularly in the microbial size range (Platt, this volume). In general, materials and energy tend to flow from the small to the large end of the size-spectrum, but there are notable exceptions, such as the reverse flow of the "microbial loop" (Azam et al. 1983). Studying the microbial loop in terms of the turnover times which characterize its elements should prove fruitful. Parasites which are smaller than their "prey" provide another example of flows down the size spectrum (Cousins, this volume). While the most detailed p.s.d.'s have been documented in soft-substrate benthic communities (Platt, this volume), wider investigation on hard substrates is required to test the generality of the allometric hypothesis.

Little is known about whether the allometric hypothesis might hold for nonliving components (e.g., DOM, POM, marine snow, etc.). The chemical composition and structure of different non-living components may to a large extent determine their turnover time in biogeochemical cycles, and it seems likely that most smaller components would turn over faster than do larger ones. Physical mechanisms (e.g., sedimentation, resuspension, and fractionation) may also cause the retention time of the nonliving components of a system to vary with size. Thus, living and nonliving components should be distinguished, the relationships between their sizes and turnover times should be cast in allometric form, and the two allometries compared. It may be that turnover time and size (of both living and nonliving particles) are related to some still undefined measure of

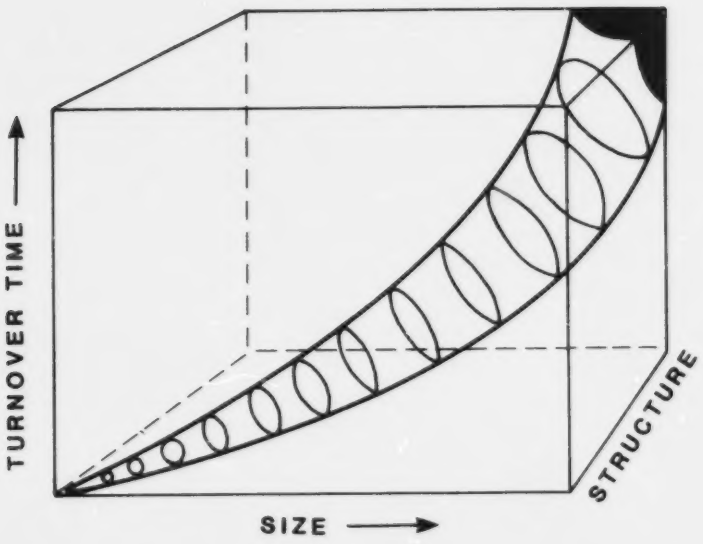


FIG. 1. Hypothetical three-dimensional relationship between "structure," turnover time, and size of particles. It is suggested that "structure" is related either to the complexity or to the energy required to disrupt the structure, whether it be a molecule, organism, or possibly even an ecosystem. At present the relationship is too hypothetical to be able to give scales for the axes.

structure, which might be quantified by the energy required to disrupt the structure (Fig. 1).

Succession in marine ecosystems may also be interpreted in terms of structure, size, and time. Stable deep-sea environments allow complex community structures to develop, whereas rocky shore communities have less complex structure due to their periodic disruption by physical forces. Plankton communities are disrupted even more frequently by the auxiliary energy provided by advective mixing, and they consist mainly of very small organisms having a rapid turnover rate (Margalef, this volume). The small size of most planktonic organisms is probably due to their need to rapidly respond to pulses of auxiliary energy. Thus, the frequency of the physical forcing events probably determines the turnover time, which in turn determines the size. Structured systems operating on longer time scales (e.g., deep-sea benthic systems, hydrothermal vents) are closer to steady-state, tend to be internally self-regulated, and undergo slow successional changes, whereas less structured systems (e.g., plankton) dominated by external forces are prone to change frequently and rapidly (Sanders 1985).

Time and Space Scales

Although observations of nature should be made on large spatial scales in order to define system boundaries, they are unlikely to provide information on the rates of processes. Ecosystem-level spatial scales tend to be large, but the characteristic rates of different component processes vary greatly and must be measured at appropriate temporal scales. Thus, a whole ecosystem or whale population might require measurements ranging over

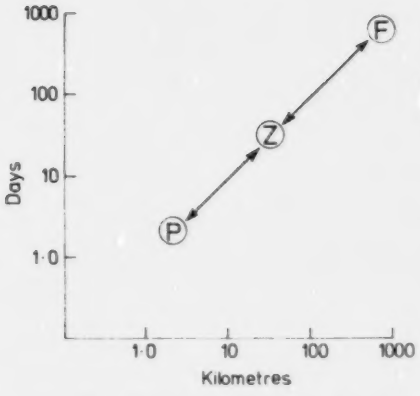


FIG. 2. Simplified representation of typical temporal and spatial scales associated with phytoplankton (P), herbivorous zooplankton (Z), and pelagic fish (P). (From Steele 1978).

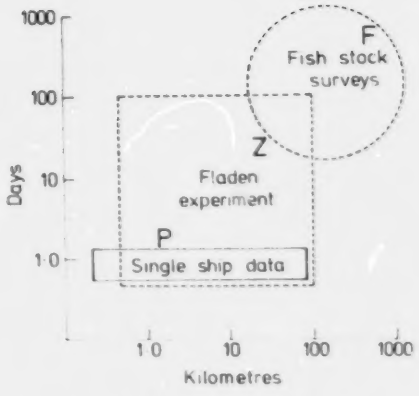


FIG. 4. An indication of the temporal and spatial scales covered by various types of sampling programs; the Fladen experiment was a multi-ship exercise carried out in the North Sea. (From Steele 1978).

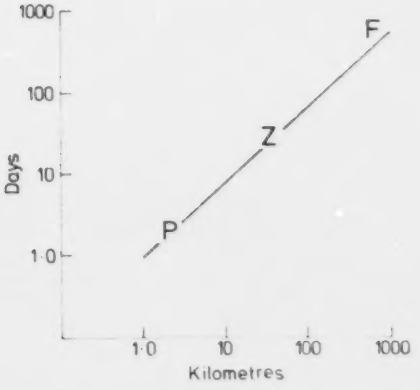


FIG. 3. Comparison of mixing scales with the trophic scales of Fig. 2. (From Steele 1978).

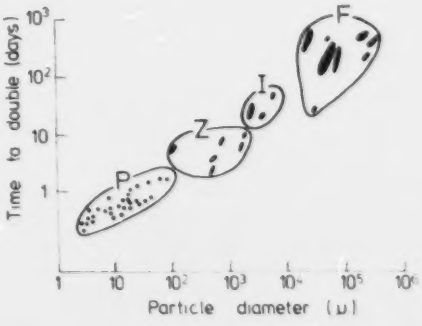


FIG. 5. The relationship between doubling time and particle size from data for phytoplankton (P), zooplankton (Z), invertebrate carnivores or omnivores (I), and fish (F). (From Sheldon et al. 1972).

several hundreds of kilometers and occurring over decades, whereas bacterial turnover rates must be measured over hours in liter or so incubations.

Figures 2–5 show some of the interactions between spatial and temporal scales as they affect biological oceanographic research (Steele 1978). The interaction between spatial scales and turnover times is depicted in Fig. 2 and 3. They also indicate the relationship between physical and biological processes (see also Legendre and Demers 1984; Legendre et al., this volume). Figure 4 shows the temporal and spatial scales covered by different types of sampling programs, and Fig. 5 shows the relationship between particle size and doubling time in marine systems. These relations indicate that once the system of interest

has been defined, there are definite constraints on experimental and survey design (see Platt et al. 1981). It is as important to obtain time series of appropriate frequency as it is to obtain synoptic data over an appropriate spatial scale. To complicate matters further, to measure the rates of processes occurring within ecological systems it is necessary to conduct experiments at temporal and spatial scales appropriate to each component, as suggested by Fig. 5. A general relationship between scales in nature and the appropriate temporal and spatial dimensions for microcosm and mesocosm experiments has yet to be fully resolved.

Recommendations for Hypothesis Testing

The group found that the most promising and readily testable hypotheses presented at the conference were:

1) Particle size distributions can be used to characterize ecosystems and how they react to perturbations. Characteristic p.s.d.'s vary according to whether the system is oligotrophic or eutrophic, pulsed or non-pulsed, and to which successional stage it is in.

2) The spectra of biological turnover rates are strongly correlated to the power spectra of physical forces impinging on the biota.

3) Measures of total system throughput of energy or material, the diversity of those flows as well as their throughput/recycling ratios can all be used to characterize ecosystems.

Implications for Sampling and Analytical Strategies

1) The allometric relationships that exist between particle sizes and turnover rates, amplify the utility of measuring particle size distributions. Measurements of p.s.d.'s in different types of aquatic ecosystem can be used to test Hypothesis 1. The living and nonliving components of particle-size spectra should be distinguished, and spectra should span as wide a range of sizes as possible in view of the importance of small particles to energy and material flows in ecosystems.

2) If Hypothesis 2 is valid, then information should be collected at the appropriate temporal and spatial scales that characterize different living, as well as nonliving, components of an ecosystem. This is imperative if the properties of the whole system are to be adequately understood. The frequencies and scales of physical forcing events and processes are likely to be of overriding importance in deciding the scales at which a system is to be observed. Appropriate spacing of synoptic measurements is as important as the correct interval in time series observations.

3) The implications of scale are also important to the design of experimental studies and should be used to determine the size (volume) of microcosm and mesocosm enclosures as well as the duration of the experiments. However, a general relationship between scales in nature and those appropriate for experiments has not yet been developed.

4) The need to measure material and energy flows in addition to biomasses is emphasized. Flows are one measure of the dynamics of system processes and permit the testing of Hypothesis 3

5) Input–output, flow, and network analyses are an extremely useful family of tools for analyzing and understanding networks of material and energy flows. These methodologies are germane to testing Hypothesis 3. To date flow data have been assembled for very few ecosystems, and the potential of recently developed network analysis techniques should be explored.

6) Properly used, modelling and simulation techniques are important tools for understanding ecosystem properties and should be used iteratively in both the planning and execution of ecosystem research programs. Models have a role to play in describing, clarifying and testing all three hypotheses outlined above.

References

AZAM, F., T. FENCHEL, J. G. FIELD, J. S. GRAY, L.-A. MEYER-REIL, AND F. THINGSTAD. 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10: 257–263.

CONRAD, M. 1983. Adaptability: the significance of variability from molecule to ecosystem. Plenum Press, New York, NY.

DAAN, N. 1980. A review of the replacement of depleted stocks by other species and the mechanisms underlying such replacement. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 177: 405–421.

LEGENDRE, L., AND S. DEMERS. 1984. Towards dynamic biological oceanography and limnology. Can. J. Fish. Aquat. Sci. 41: 2–19.

MANDELBROT, B. B. 1983. The fractal geometry of nature. Freeman, New York, NY.

MANN, K. H. 1982. Ecology of coastal waters: a systems approach. Blackwell, Oxford.

MARGALEF, R., AND E. GUTIERREZ. 1983. How to introduce connectance in the frame of an expression for diversity. Am. Nat. 121: 601–607.

PETERS, R. H. 1983. The ecological implications of body size. Cambridge Univ. Press, Cambridge, U.K.

PLATT, T., K. H. MANN, AND R. E. ULANOWICZ [ED.]. 1981. Mathematical models in biological oceanography. UNESCO Press, Paris.

SANDERS, H. L. 1985. Sulfide research in the marine environment and its implications for the upwelling system. Proc. Int. Symp. on upwelling areas off Africa, Barcelona. Investigacion Pesquera. (In press)

SHELDON, R. W., A. PRAKASH, AND W. H. SUTCLIFFE. 1972. The size distribution of particles in the ocean. Limnol. Oceanogr. 17: 327–340.

STEELE, J. H. 1978. Some comments on plankton patches. *In* J. H. Steele [ed.] Spatial pattern in plankton communities. Plenum Press, New York, NY.

ULANOWICZ, R. E. 1980. An hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

WEBSTER, J. R. 1979. Hierarchical organizations of ecosystems. *In* E. Halfon [ed.] Theoretical systems ecology. Academic Press, New York, NY.

III. Possible Holistic Approaches to the Study of Biological–Physical Interactions in the Oceans

L. Legendre* (Chairman), W. M. Kemp, H. Atlan, M. Conrad, M. Fréchette, P. Lane, T. Platt, G. Rodriguez, J. Tundisi, and C. S. Yentsch

Scope

A significant part of the current research in oceanography aims at explaining biological–physical interactions in terms of mechanisms and processes. Legendre and Demers (1984), for example, recognize hydrodynamics as the driving force of aquatic ecosystems — hydrodynamic variability being transferred to the living organisms through various physical, chemical, and biological mechanisms. In addition to mechanistic studies, biological oceanographers have also utilized several integrated approaches that have been introduced in recent decades in various experimental sciences. The two approaches, mechanistic and holistic, are often considered antithetical; however, a synthetic use of both methods in biological oceanography could lead to a better understanding of marine ecosystems and possibly an improved capability for prediction. The holistic methods of allometry, irreversible thermodynamics, information theory, and network analyses will be considered here in terms of their potential utility in the study of biological–physical interactions in the oceans.

As has been discussed by Legendre and Demers (1984), the marine plankton (phytoplankton, zooplankton, and fish larvae) respond not only to the amplitude but also to the frequencies of periodic physical forcings. A large part of the effort in studying biological–physical processes over the last decade has focussed on the planktonic components of the ecosystem, since plankton respond readily to changes in the physical environment. However, there is now a need, in assessing the responses of marine ecosystems to perturbations (at various amplitudes and frequencies) of the physical environment, to extend the study of biological–physical interactions to the nekton and to the benthos as well.

Particle Size and Allometry

One obviously holistic biological property to be measured in the aquatic environment is embodied in the particle size spectrum. Particle size measurements are presently limited to a very narrow range (ca. 2 to 250 μm), as far as automatic particle counting is concerned, and only such recent developments as the flow cytometer (Yentsch et al. 1983) or a sea-going counterpart (Olson et al. 1983) make it possible to discriminate automatically between living and nonliving particles in this limited size range. Various microcosm experiments have suggested that stresses externally imposed on pelagic ecosystems can lead to shifts in the size structure of marine food-chains (e.g., O'Connors et al. 1978), which may affect the amount and the quality of secondary production (Silvert and Platt 1978; Greve and Parsons 1977; Landry 1977). Process studies have shown (see Legendre and Demers 1984) that biological and physical state properties are related only at very large spatial and/or temporal scales, or at very sharp transitions. At smaller spatiotemporal scales, rate variables for both the physical forcing and the biological responses must be considered. It can therefore be inferred that holistic rate measurements will be required

*GIROQ, Département de biologie, Université Laval, Québec (Québec) G1K 7P4.

when studying the responses of the particulate ecosystem to the frequencies of the physical perturbations.

Allometric criteria (e.g., relating size to generation time) can explain the presence or absence of certain size organisms in provinces of the sea that are characterized by different spectra of physical forces. This perspective was used by Walsh (1976) in his interpretation of upwelling ecosystems off the west coast of the Americas. According to the allometric approach (see the papers by Calder, Cousins, and Platt, this volume), turnover rate spectra should be easy to derive from particle size spectra, since turnover rate is generally a monotonic function of particle size. One possible conclusion of the allometric approach, for phytoplankton growing in a stable nutrient-limited environment, is that the smaller cells with high turnover rate should out-compete the larger cells with lower turnover rate. This is partially the case, as the smaller phytoplankton cells are proportionally more abundant in the oligotrophic oceanic waters, than in nearshore eutrophic waters (Teixeira and Tundisi 1967; Malone 1980). However, even in oligotrophic environments, the larger forms are not completely eliminated, perhaps due to differential grazing (Silvert and Platt 1980). This also could be interpreted as a consequence of fluctuations in the physical environment, which would tend to maintain a full spectrum of algal cells (including the larger ones) by locally distorting the allometric relationship between size and turnover rate. A spatiotemporal mosaic of different turnover rate spectra should therefore be maintained by the environmental fluctuations, which is consistent with the contemporaneous disequilibrium hypothesis (Richerson et al. 1970) and also with the model of Kemp and Mitsch (1979). If this hypothesis is correct, allometric relationships would then emerge only when different samples from the spatio-temporal mosaic are pooled together.

This holistic approach raises several sampling problems. The first one is that ecosystem and physical rate properties should be measured simultaneously. The frequency of sampling the various size fractions depends on both the turnover rates of the organisms and the actual frequencies of the physical perturbations. Another problem is that, in any given sample, the estimation error increases with the size of the organisms, as the larger ones are generally less abundant. On the other hand, the estimation of particle sizes in a sample is probably less representative of the parental population for the smaller organisms, since the distribution of smaller size particles tends to be controlled by the heterogeneities in the physical environment.

Interfaces as Sites of Biological–Physical Interactions

Biological–physical interactions in the oceans mediate the transfers of energy from the abiotic environment to the living organisms. The input of auxiliary mechanical energy is known to enhance energy transfers in the marine ecosystems (see Margalef, this volume). For example, Legendre (1981) has shown that the phytoplankton production potential of an aquatic ecosystem is a function of the frequency at which it alternately stabilizes and destabilizes.

Important characteristics of the physical structure of the pelagic environment are its various interfaces: the ice–water interface (epontic microflora), pycnoclines (deep chlorophyll maxima), tidal fronts (frontal chlorophyll maxima), and so forth. The benthos is also defined by the water–sediment interface, but there are other boundaries of interest on the sea floor such as the fringe of calcareous reefs and the redox discontinuity layer in marine sediments. The interfaces have been recognized for years by microbiologists as preferred biotic habitats (Marshall 1976), and they support relatively high metabolic activity (e.g., Legendre 1981; Kemp et al. 1982). At interfaces such as the sediment surface, there is a sharp gradient of kinetic energy which enables organisms to utilize the power of fluid

motion for their mechanical work (e.g., for pumping dissolved or particulate substances, Nixon et al. 1971). Sharp gradients in heat or chemical concentrations across these interfaces create potentials which facilitate the transport of food and wastes via convective and diffusive mechanisms, thus stimulating metabolic processes. These problems could probably be investigated holistically by using the formalism provided by irreversible thermodynamics (see the papers by Allen and Zotin, this volume).

At some interfaces, biological activity can modify the physical environment in such a way as to enhance the growth and/or survival of ecosystem components. For instance, Lewis et al. (1983) did show that increased absorption of solar energy on the upper slope of the deep chlorophyll maxima may affect vertical convective circulation, thereby mixing phytoplankton cells upward to more favorable light conditions. Some benthic communities (coral reefs) effectively increase the tidal currents (and associated kinetic energy available for mechanical work) in overlying water, by decreasing water depth above their growing calcareous mounds (Odum and Odum 1955). The stirring effect of such formations can induce the upwelling of nutrient-rich water (Simpson et al. 1982). Sea-grasses and other macrophytes can improve their light environments by enhancing sediment deposition and reducing resuspension (Ginsburg and Lowenstam 1958; Ward et al. 1984). There are numerous other examples of such positive feedbacks, which may constitute a generic set of ecosystem properties worth investigating in the light of thermodynamics.

Both at interfaces and in the water column, patchiness is a general characteristic of ecosystem organization in the oceans (phytoplankton patches, zooplankton clouds, fish schools, benthic aggregates, etc.). The mechanistic explanations of phytoplankton patchiness invoke the interplay between hydrodynamics and phytoplankton growth (Skellam 1951; Kierstead and Slobodkin 1953; Denman and Platt 1976; Denman et al. 1977), as well as nutrient limitation (Wroblewski et al. 1975) and zooplankton grazing (Platt and Denman 1975; Wroblewski et al. 1975; Riley 1976). For zooplankton clouds, behavioral interactions also play a role (Lee and McAlice 1979; Haury and Wiebe 1982), as they also do in other animal aggregrates. These various mechanisms are perhaps amenable to a more general holistic approach, as proposed in the next paragraph.

Coherence Between Physical and Biological Events

Interactions between patches of organisms and the physical environment are usually defined using cross-correlations, or coherence and phase spectra. Cross-correlations can be expressed in terms of conditional entropies, using the formalism of information theory (see the papers of Atlan and Conrad, this volume). The main advantage of the information theoretical approach is that it provides a mathematical framework for dealing with the relationships between physical features of the environment and observed biological structures. Information measures may be used to characterize the uncertainty of the environment, while the conditional entropy structure can be used to analyse how environmental disturbances are propagated through the community. The patchiness of the community makes a major contribution to its conditional entropy structure. It may therefore be assumed that simultaneous measurements of environmental uncertainty (still to be defined in practical operational terms) and the conditional entropy structure of the ecosystem could contribute to predicting the responses of patchy communities to external perturbations, and to forecasting those responses of higher trophic levels to various changes in the patchy lower levels.

A way to investigate the responses of communities to environmental perturbations is to measure their adaptability (Conrad 1983). Up to now, adaptability processes have been analyzed according to the levels of biological organization. Another approach, for marine

ecosystems, would be to consider adaptability as a function of both the size of the organisms (allometry) and the frequencies of the environmental forcing. How to actually measure adaptability in marine ecosystems is a critical problem. Indices of environmental uncertainty serve as the measure of required adaptability. These should be related in a definite way to indices of diversity and patchiness, and to genetic and physiological variabilities of the constituent taxa (Conrad, this volume). Adaptability may therefore be another means by which to analyze holistic relationships between changes in the physical environment and changes in the ecosystem.

Network Techniques

The various techniques for the analysis and simulation of marine ecological networks represent a set of potentially powerful tools for studying holistic properties of biological–physical interactions. None of these methods seems to be inherently more or less holistic than the others, the insight gained about marine ecosystems being a function of how a particular method is used. Static techniques such as energy-circuit language (Odum 1967, 1971) or input–output analyses (see the papers by Fasham, Hannon, and Patten, this volume) offer the ability for the a posteriori analysis of network structure. The network structure is an holistic ecosystem property, the variations of which could be studied by reference to the physical forcing. Techniques that rely on the topology of ecological networks, such as loop analysis (e.g., Lane and Collins 1985) or dynamic network analysis (Mikulecky, this volume), as well as the network information indices defined by Ulanowicz (1980), can be used to assess the qualitative effects of perturbations on ecosystems. When actual flows are available, such techniques as dynamic network analysis also allow quantitative simulations. Conceptually aggregated, less-mechanistic simulation models offer an approach to identifying and understanding emergent ecosystems responses to physical conditions (e.g., O'Neill 1976) or perturbations at various frequencies (Odum 1983). Ultimately, the utility of these techniques is limited by the availability of qualitative, or quantitative, empirical descriptions of marine ecosystems.

Sampling Design

A final cautionary remark concerns the selection of biological and physical variables to be measured when studying marine ecosystems. If empirical evidence and/or theoretical analysis show that the sampling design will bias the estimation of a given variable, either this variable should not be sampled, or the sampling should be re-designed. Too often, variables which are known to change rapidly or to be distributed heterogeneously are estimated from widely spaced samples, thus leading to biased estimates of both the mean and the variance and thus to questionable predictions. We therefore suggest intensive sampling programs of holistic properties, on spatio-temporal scales that are compatible with both the turnover rates of the organisms and the hydrodynamic rates.

References

CONRAD, M. 1983. Adaptability: the significance of variability from molecule to ecosystem. Plenum Press, New York, NY.

DENMAN, K. L., A. OKUBO, AND T. PLATT. 1977. The chlorophyll fluctuation spectrum in the sea. Limnol. Oceanogr. 22: 1033–1038.

DENMAN, K. L., AND T. PLATT. 1976. The variance spectrum of phytoplankton in a turbulent ocean. J. Mar. Res. 34: 593–601.

GINSBURG, R. N., AND H. A. LOWENSTAM. 1958. The influence of marine bottom communities on the depositional environment of sediments. J. Geol. 66: 310–318.

GREVE, W., AND T. R. PARSONS. 1977. Photosynthesis and fish production: Hypothetical effects of climatic change and pollution. Helgol. Wiss. Meeresunters. 30: 666–672.

HAURY, L. R., AND P. H. WIEBE. 1982. Fine-scale multi-species aggregations of oceanic zooplankton. Deep-Sea Res. 29: 915–921.

KEMP, W. M., AND W. J. MITSCH. 1979. Turbulence and phytoplankton diversity: a general model for the "paradox of plankton". Ecol. Model 7: 201–222.

KEMP, W. M., R. L. WETZEL, W. R. BOYNTON, C. F. D'ELIA, AND J. C. STEVENSON. 1982. Nitrogen cycling and estuarine interfaces, p. 209–230. *In* V. S. Kennedy [ed.] Estuarine comparisons. Academic, New York, NY.

KIERSTEAD, H., AND L. B. SLOBODKIN. 1953. The size of water masses containing plankton blooms. J. Mar. Res. 12: 141–147.

LANDRY, M. R. 1977. A review of important concepts in the trophic organization of pelagic ecosystems. Helgol. Wiss. Meeresunters. 30: 8–17.

LANE, P., AND T. M. COLLINS. 1985. Food web models of a marine plankton community: an experimental approach. J. Exp. Mar. Biol. Ecol. (In press)

LEE, W. Y., AND B. J. MCALICE. 1979. Sampling variability of marine zooplankton in a tidal estuary. Estuarine Coastal Mar. Sci. 8: 565–582.

LEGENDRE, L. 1981. Hydrodynamic control of marine phytoplankton production: the paradox of stability, p. 191–207. *In* J. C. J. Nihoul [ed.] Ecohydrodynamics. Elsevier, Amsterdam.

LEGENDRE, L., AND S. DEMERS. 1984. Towards dynamic biological oceanography and limnology. Can. J. Fish Aquat. Sci. 41: 2–19.

LEWIS, M. R., J. J. CULLEN, AND T. PLATT. 1983. Phytoplankton and thermal structure in the upper ocean: consequences of nonuniformity in chlorophyll profile. J. Geophys. Res. (C) 88: 2565–2570.

MALONE, T. C. 1980. Algal size, p. 433–463 (chapter 12). *In* I. Morris [ed.] The physiological ecology of phytoplankton. Blackwell, Oxford.

MARSHALL, K. C. 1976. Interfaces in microbial ecology. Harvard Univ. Press, Cambridge, MA. 156 p.

NIXON, S. W., C. A. OVIATT, C. ROGERS, AND K. TAYLOR. 1971. Mass and metabolism of a mussel bed. Oecologia 8: 21–30.

O'CONNORS, H. B., C. F. WURSTER, C. D. POWERS, D. C. BIGGS, AND R. G. ROWLAND. 1978. Polychlorinated biphenyls may alter marine trophic pathways by reducing phytoplankton size and production. Science 201: 737–739.

ODUM, H. T. 1967. Biological circuits and the marine systems of Texas, p. 99–157. *In* T. A. Olson and J. J. Burgess [ed.] Pollution and marine ecology. Wiley Interscience, New York, NY.

1971. Environment, power and society. John Wiley, New York, NY.

1983. Systems ecology: an introduction. John Wiley, New York, NY. 644 p.

ODUM, H. T., AND E. P. ODUM. 1955. Trophic structure and productivity of a windward coral reef community on Eniwetok Attoll. Ecol. Monogr. 25: 292–320.

OLSON, R. J., S. L. FRANKEL, S. W. CHISOLM, AND H. M. SHAPIRO. 1983. An inexpensive flow cytometer for the analysis of fluorescence signals in phytoplankton: chlorophyll and DNA distributions. J. Exp. Mar. Biol. Ecol. 68: 129–144.

O'NEILL, R. V. 1976. Ecosystem persistence and heterotrophic regulation. Ecology 57: 1244–1253.

PLATT, T., AND K. L. DENMAN. 1975. A general equation for the mesoscale distribution of phytoplankton in the sea. Mem. Soc. R. Sci. Liege (Ser. 6) 7: 31–42.

RICHERSON, P., R. ARMSTRONG, AND C. R. GOLDMAN. 1970. Contemporaneous disequilibrium: a new hypothesis to explain the "paradox of the plankton". Proc. Natl. Acad. Sci. U.S.A. 67: 1710–1714.

RILEY, G. A. 1976. A model of plankton patchiness. Limnol. Oceanogr. 21: 873–880.

SILVERT, W., AND T. PLATT. 1978. Energy flux in the pelagic ecosystem: A time-dependent equation. Limnol. Oceanogr. 23: 813–816.

1980. Dynamic energy-flow model of the particle size distribution in pelagic ecosystems. Am. Soc. Limnol. Oceanogr. Spec. Symp. 3: 754–863.

SIMPSON, J. H., P. B. TETT, M. L. ARGOTE-ESPINOZA, A. EDWARDS, K. J. JONES, AND G. SAVIDGE. 1982. Mixing and phytoplankton growth around an island in a stratified sea. Cont. Shelf Res. 1: 15–31.

SKELLAM, J. G. 1951. Random dispersal in theoretical populations. Biometrika 78: 196–218.

TEIXEIRA, C., AND J. TUNDISI. 1967. The phytoplankton and primary production in waters of the equatorial region. Bull. Mar. Sci. 17: 884–891.

ULANOWICZ, R. E. 1980. An hypothesis on the development of natural communities. J. Theor. Biol. 85: 223–245.

WALSH, J. J. 1976. Herbivory as a factor in patterns of nutrient utilization in the sea. Limnol. Oceanogr. 21: 1–13.

WARD, L. G., W. M. KEMP, AND W. R. BOYNTON. 1984. The influence of waves and seagrass communities on suspended particulates in an estuarine embayment. Mar. Geol. 59: 85–103.

WROBLEWSKI, J. S., J. J. O'BRIEN, AND T. PLATT. 1975. On the physical and biological scales of

phytoplankton patchiness in the ocean. Mem. Soc. R. Sci. Liège (Ser. 6) 7: 43–57.

YENTSCH, C. M., P. K. HORAN, K. MUIRHEAD, Q. DORTCH, E. HAUGEN, L. LEGENDRE, L. S. MURPHY, M. J. PERRY, D. A. PHINNEY, S. A. POMPONI, R. W. SPINRAD, M. WOOD, C. S. YENTSCH, AND B. J. ZAHURANEC. 1983. Flow cytometry and cell sorting: a technique for analysis and sorting of aquatic particles. Limnol. Oceanogr. 28: 1275–1280.

IV. Technological Developments to Implement Theory into Biological Oceanography

K. Denman* (Chairman), W. Calder, C. Davis (Rapporteur), S. Demers, M. Estrada, M. Lewis, D. Smith, P. Wangersky, C. Yentsch, and A. Zotin

Introduction

The recent concepts in theoretical ecology presented at this conference, while thought-provoking, are for the most part still developing. As a consequence, many of the ideas have not yet been formulated in terms of observable or measurable variables. Their application to marine ecosystems will be a gradual process, and we are not currently in a position to outline a definitive experiment or to design and build specific new instruments. We believe that the next step, which has been started by Platt (this volume) and Fasham (this volume), is to apply these principles *now* to existing whole ecosystem data sets to see how these new approaches might provide insights beyond our present understanding of marine ecosystem dynamics. In this report we will endeavor to state some criteria to be considered in any field study of whole marine ecosystems, to identify several obvious gaps in our knowledge based on the material presented at this conference, and to outline some necessary technological developments that ought to be recommended and encouraged by SCOR Working Group 73.

Criteria for the Study of Marine Ecosystems

1) The marine ecosystem is *strongly forced* by the physical environment at many temporal and spatial scales. We cannot understand or begin to predict the time dependent behavior of the marine ecosystem without also understanding the behavior of the ocean–atmosphere system and how the marine ecosystem responds to it. The ocean and the atmosphere are both turbulent fluids, both far from any type of equilibrium and thereby foiling any simple understanding of their dynamics on our part. In addition, if we want to predict future behavior of the marine ecosystem, then obviously we must also predict and specify the behavior of the primary forcing functions, i.e., the ocean and the atmosphere.

2) The importance of synoptic measurements of *rate processes* as adjuncts to the more traditional biomass measurements should be emphasized. In the absence of data delimiting the scales at which a quantity may vary, we should try to take measurements as closely spaced in time and space as is feasible, e.g., try to resolve both diel and annual variations. Records of observations should be complemented by data from controlled experimental manipulations in the search for causal relationships, e.g., denuding certain areas or components of a benthic ecosystem where such experiments are possible. Since natural fluctuations are so prevalent, we should strive to use them to analyze the response of marine ecosystems to perturbations.

3) *Undersampling* the variability of components and processes will probably always occur in studies of pelagic marine ecosystems, and inference from incomplete data sets will therefore be necessary. We will be required to compromise between spatial and temporal resolutions (in fact, one will often be confounded by the other), between measuring biomasses and rates, and between understanding internal processes and considering

*Institute of Ocean Sciences, P.O. Box 6000, 9860 W. Saanich Rd., Sidney, B.C. V8L 4B2.

only inputs and outputs. Those designing sampling programs will have to consider which temporal and spatial scales are of greatest interest (i.e., the sampling 'window'), and we will have to assimilate and incorporate the different types of information obtained into our models and analyses.

Observational Gaps Suggested by the Conference Presentations

The shape of particle or organism size frequency distributions may be predicted from allometric relationships (Platt, this volume). The allometric approach, developed for terrestrial mammals, may not apply generally to the diverse organisms found in marine ecosystems. The only consistent evidence that allometry is pertinent to marine biology resides in a very few measurements taken in the "steady-state" mid-ocean gyres. The robustness of the slope of allometric relations must be ensured through careful observations made in a wide range of environments and over a broad range of size classes. (For example, whales in a given ocean gyre must be counted for at least several generations before an average census is meaningful.)

Even if the slope of the distribution could be accurately determined, knowing the absolute number density in any size class would require an accurate estimate of the net flow of biomass from very small organisms (with equivalent diameter of $< \sim 0.2$ mm, where photosynthesis occurs) to the larger organisms. Important flows occur at these microscales which need to be assessed: in particular, a significant fraction of the community respiration occurs here, and significant amounts of mass and energy flow back from larger (biogenic detritus and fecal materials) to smaller (bacteria) particles. To evaluate these flows, we must rely on measurement techniques that *identify various types of particles within the same size class*. The ideas presented on flow analysis underscore this need: for example, Hannon (this volume) introduced the concept of the unit 'value' of a commodity, which as we understand it requires knowing the identity of various particles. Fasham's (this volume) flow analysis indicated that as much as 40% of the primary production may be respired in the microbial loop and that 75% of the nitrogen taken up by phytoplankton may be recycled. Patten (this volume) showed the importance of storage and residence times in multiply-connected food webs, delays which should be incorporated into Fasham's analysis.

Aggregates of 'marine snow' are particularly interesting components of particle size distributions. To study the material flows along the particle size spectrum that result from growth or the grazing of bacteria by micro-organisms living on particles, we may consider the actions of individual organisms. However, to quantify properly the material flow due to grazing by larger organisms, we must consider how whole aggregates interact. Since most rapid organism counting systems employ pumping or concentration by nets, aggregates are damaged before and during measurement; thus we identify a need to develop *nondestructive high data rate particle samplers*, such as in situ optical devices.

Finally, thermodynamics, as applied to organisms rather than to networks (e.g., Zotin, this volume), suggests a need for more measurements pertaining to organisms in the microbial loop. To be more specific, it would appear useful to couple measurements of oxygen flows from microrespirometry and heat flows from microcalorimetry.

Sampling Platforms

We must coordinate our sampling of the biota with observations of the physical environment if we are to understand and to predict the coupling of biological dynamics with physical processes. We advocate *matched sampling protocols*, whereby biological and

physical variables are measured simultaneously at the same sampling rate and (preferably) stored together on the same medium. Biologists have been hampered by the lack of *unattended platforms*, both drifting and moored; physical oceanographers have been using such platforms for nearly two decades to make observations over temporal and spatial intervals that are unfeasible with shipboard operations. Recently, fluorometers have been moored and left unattended (Whitledge and Wirick 1983). The scope of such automated measurements must be expanded to include many biotic and abiotic variables, and in situ computers should be incorporated for purposes of data reduction, error checking and storage (Smith and Horner 1982).

Such automated systems can resolve most temporal variability and, if towed or densely arrayed, some spatial variability. However, *remote sensing* from satellites or airplanes will probably be necessary to encompass large-scale horizontal patterns synoptically. Currently, phytoplankton pigment and sediments can be measured from space with multicolor scanners, and sea surface temperature with infrared sensors (e.g., Smith and Baker 1982; Borstad et al. 1982). These sensors are mounted on different satellites, and the one satellite with an ocean color sensor is now operating long past its projected lifetime. Replacing and consolidating these satellites should be given top priority. The feasibility of including active sensing techniques, such as Raman backscattering of laser light, into the next generation of oceanographic satellites should also be investigated.

Sensor Systems

1) Particle counters — Coulter counters have long been used to assess particle size distributions in the sea. Another very exciting system (developed initially for medical use) is the *flow cytometer and sorter*, which can be used to measure simultaneously, at very rapid rates, multiple properties of individual cells and particles. We can exploit lasers to excite specific fluorescent frequencies of pigments and stains to serve as signals to identify and quantify subpopulations of cells and particles in the 1 to 150 μm size range, at rates that exceed 1000 cells/s (Yentsch et al. 1983). At larger diameters (0.3−4 mm), *zooplankton counters* with pumps and towed in situ vehicles can now automatically and continuously estimate organism diameter, length and volume (Herman and Mitchell 1981; Mackas et al. 1981). These sensors are routinely deployed alongside nets, which are used to collect calibration samples, but they would be better calibrated by simultaneous data from nondestructive samplers.

2) Nondestructive samplers — The classic nondestructive sampler is the camera, and high quality photographs have been obtained recently by *shadowgraph* techniques (Ortner et al. 1981). To use photographs effectively in conjunction with continuous techniques, we need further development of fast *computer image analysis* techniques. Other optical properties of the sea (e.g., absorption, scattering and attenuance) that can be measured continuously are strongly influenced by the quantity and type (size and pigment composition) of biogenic and abiogenic material present. Consequently, measurement of these optical properties permits the use of inversion methods to estimate biological properties of the marine ecosystem. Both *passive and active optical sensors* are available; their power requirements have been reduced to the point that unattended operation for up to 30 days is now possible. Larger organisms can be assessed nondestructively by acoustical means. *Multiple frequency acoustical systems* should provide data that can be analysed to yield size distributions of the 'micronekton' (e.g., Pieper and Holliday 1985; Greenlaw and Johnson 1983).

3) Automatic chemical sensors — We must develop more sensors for use on unattended platforms than the few currently available. Two types of sensors are possible: *electrodes*

for specific ions or elements (e.g., oxygen), and *flow injection spectrophotometric systems*, where certain reagents are periodically injected into samples, which are thereafter scanned by a spectrophotometer. These systems will be relatively costly, but these costs will not even equal the resources required by some of the global programs in physical oceanography now being planned (e.g., World Ocean Circulation Experiment [WOCE]; Webster, 1984).

4) Rates of growth, respiration, and adaptation — Rate processes vary on a number of ecologically important scales, but our ability to sample these rates at high frequencies is severely limited. We must focus our efforts on developing *instantaneous measurement techniques* that permit direct or indirect estimation of rates without requiring incubation over extended intervals of time. For example, the 'photosynthetron' (Lewis and Smith 1983) permits the rapid determination of photosynthesis-irradiance curves. *Fluorometric techniques* may also provide suitable *proxy* variables for the processes of interest. Recently, an instantaneous estimate of an index of phytoplankton photoadaptation (Demers et al. 1985) was provided by the continuous measurement of the fluorescence kinetics induced in the phytoplankton by exposing them to bright light. Work is also in progress to develop millisecond double flash techniques that will allow one to infer instantaneous photosynthetic rates (Falkowski 1983). Metabolic rates are now being measured directly in microcosms by means of a *microcalorimetry* apparatus (Tournie and Lasserre 1984; Lasserre and Tournie 1984). Work is underway to develop a flow-through system to make these measurements in situ. These techniques are less applicable to measure metabolic rates of larger organisms, but video measurements of *swimming activity* are currently being used to infer rates of feeding by zooplankton.

Concluding Remarks

The Working Group felt that none of the theoretical ideas presented at the conference compel the design and construction of totally new instrumentation. Rather, we have recommended directions for the development of existing instrumentation and techniques into new forms that will yield the observations most needed to test and apply these theoretical ideas. We do not feel that this pragmatic view will limit the application of theory, but rather that it will hasten the testing and incorporation of new theoretical concepts into marine ecology. Finally, we feel that the testing and evaluation of new ideas, new techniques, and new measurement tools should be an ongoing, iterative process, not a "one shot" affair.

References

BORSTAD, G., R. BROWN, D. TRUAX, T. MULLIGAN, AND J. GOWER. 1982. Remote sensing techniques for fisheries oceanography: examples from British Columbia. NAFO Sci. Coun. Studies 4: 69–76.

DEMERS, S., J. THERRIAULT, L. LEGENDRE, AND J. NEVEUX. 1985. Continuous *in situ* measurement of the index of phytoplankton photoadaptation (I_k), using an *in vivo* fluorescence ratio. Limnol. Oceanogr. (In press)

FALKOWSKI, P. 1983. Non-destructive estimation of ongoing photosynthetic rates using a delayed double flash fluorescence technique. Abstract. EOS: Trans. Amer. Geophy. Un. 64: 1047.

GREENLAW, C., AND R. JOHNSON. 1983. Multiple-frequency acoustical estimation. Biol. Oceanogr. 2: 227–252.

HERMAN, A., AND M. MITCHELL. 1981. Counting and identifying copepod species with an *in situ* electronic zooplankton counter. Deep-Sea Res. 28: 739–755.

LASSERRE, P., AND T. TOURNIE. 1984. Use of microcalorimetry for the characterization of marine metabolic activity at the water–sediment interface. J. Exp. Mar. Biol. Ecol. 74: 123–139.

LEWIS, M., AND J. C. SMITH. 1983. A small volume, short-incubation-time method for measurement of photosynthesis as a function of incident irra-

diance. Mar. Ecol. Prog. Ser. 13: 99–102.

MACKAS, D., T. CURRAN, AND D. SLOAN. 1981. An electronic zooplankton counting and sizing system. MTS-IEEE, Oceans 1981 Conference Record, 783–786.

ORTNER, P., L. HILL, AND H. EDGERTON. 1981. *In-situ* silhouette photography of Gulf Stream zooplankton. Deep-Sea Res. 28: 1569–1576.

PIEPER, R. E., AND D. V. HOLLIDAY. 1985. Acoustical measurements of zooplankton distributions in the sea. J. Cons. (In press)

SMITH, D. F., AND S. M. J. HORNER. 1982. Laboratory and field measurements of aquatic productivity made by a minicomputer employing a dual oxygen electrode system. Mar. Biol. 72: 53–60.

SMITH, R. C., AND K. S. BAKER. 1982. Oceanic chlorophyll concentrations as determined by satellite (Nimbus-7 Coastal Zone Color Scanner). Mar. Biol. 66: 269–279.

TOURNIE, T., AND P. LASSERRE. 1984. Microcalorimetric characterization of seasonal metabolic trends in marine microcosms. J. Exp. Mar. Biol. Ecol. 74: 111–121.

WEBSTER, F. 1984. An Ocean Climate Research Strategy. National Academy Press, Washington, D.C. 66 p.

WHITLEDGE, T., AND C. WIRICK. 1983. Observations of chlorophyll concentrations off Long Island from a moored *in situ* fluorometer. Deep-Sea Res. 30: 297–309.

YENTSCH, C. M., P. HORAN, K. MUIRHEAD, Q. DORTCH, E. HAUGEN, L. LEGENDRE, L. MURPHY, M. PERRY, D. PHINNEY, S. POMPONI, R. SPINRAD, M. WOOD, C. S. YENTSCH, AND B. ZAHURANEC. 1983. Flow cytometry and cell sorting: a technique for analysis and sorting of aquatic particles. Limnol. Oceanogr. 28: 1275–1280.

V. The Design of Large-Scale Cooperative Experiments

P. A. BERNAL* (CHAIRMAN), S. COUSINS, R. HAEDRICH, B. J. HANNON, H. HIRATA, C. JOIRIS, P. LASERRE, B. C. PATTEN, S. SATHYENDRANATH, AND V. SMETACEK

Scope

Following the terms of reference from SCOR WG-73, the Group reviewed the existing alternative approaches to studying the global, holistic properties of marine ecosystems. It was recognized that many single-process components of ecosystems are now considerably better understood, so that emphasis was trained on critically assessing the current status of describing coupled-processes as they occur in marine ecosystems.

Recommendations

The Working Group recognized that a gulf exists between the various theories presented in the five sessions of this meeting and their practical implementation in Biological Oceanography. Our discussions centered on the means for bridging this gap. Conceptual modelling at different scales was considered to be a very useful strategy for discovering common interests and for increasing interaction between theoreticians and biological oceanographers. Furthermore, it was emphasized that the pursuit of ecosystems research requires copious resources and hence must be effectively organized. As a first step to these ends, a *multidisciplinary conceptual modelling exercise* might contribute quite effectively by yielding models testable through both observation and experiment.

A clear distinction was made between large-scale observational programs and true ecosystem level experiments. Strictly speaking, "experiments" imply manipulation under controlled conditions. At present, such manipulations are not feasible in the field, and microcosms or mesocosms provide the only substitute for this ideal. In view of the difficulties inherent in extrapolating the results from experiments run in enclosures to events in the natural world, it was felt that incorporating small-scale experiments into the framework of large-scale observational programs would be the most profitable compromise now possible.

Responsiveness to physical forces is a fundamental characteristic of marine ecosystems. The impact of this "forcing" occurs at several spatial and temporal scales, all of which need to be considered. The Working Group felt that different marine ecosystems — from coastal to oceanic — subject to varying spectral patterns of forcing, should be compared. For example, the wealth of data available from coastal marine ecosystems could be more judiciously exploited by detailed intercomparisons. Some physically well-defined enclosed environments, such as fjords, coastal lagoons, and others with restricted advection, are also well suited for comparative study.

Another potentially useful criterion for delimiting the systems to be studied is to focus attention on ecosystems within a geographical area occupied by a discrete fish population. Here the information provided by the commercial catch would be of value. Furthermore, comparisons between different species or groups of fish having similar habitat requirements but occupying different areas, could provide additional insights into ecosystem dynamics. In this context, fishing activities could be considered as a periodic perturbation, the response to which could help to elucidate community dynamics.

---

*(BIOTECMAR,) Pontificia Universidad Católica de Chile, Sede Regional Talcahuano, Chile.

The Group endorsed the conclusions put forward by SCOR WG-59 regarding the importance of "measuring physiological rates for the computation of ecological fluxes," in order to gain access to the dynamics of marine ecosystems. Microcosm experiments (shipboard or laboratory) should allow the measurement of such rates.

The Group suggests that the above-outlined modelling exercise should serve as a starting point for approaching a continuum of research objectives, and that as progress is made, efforts should be switched into observational research programs designed to test specific emerging hypotheses.

Other Considerations

In reviewing the presentations of the five sessions of this conference the Group recognized the positive impact that the new concepts could have on the study of holistic properties of ecosystems and appreciated the improved methodologies now available to analyze formal constructs. For example, particle-size spectrum theory, when merged with the allometric approach, holds fresh promise for the elucidation of ecosystem structure and function. Concerning the comparative studies suggested earlier, the Group considered it worthwhile to make an effort to obtain data covering the entire size spectrum of ecosystems in specific geographical areas. Anomalous pathways, e.g., those defined by parasitic relationships, should be given special consideration. Existing data could be used to complete the size spectra, especially at the larger scales.

New modes of analysis presented at the Conference permit one to evaluate the connectedness of systems. For example, flow or environ analysis now allow one to gauge the importance of indirect effects. Network thermodynamics and information theory will gain in relevance if ecosystems can be successfully described with these tools. These methodologies need not be applied independently, but are best tried in concert or parallel.

Implementation

To implement the above recommendations a network of those scientists and institutions interested in cooperative, comparative studies should be established. The Group recommends that SCOR and IABO proceed to identify institutions and scientists interested in participating in such a program. Training programs for young scientists should be included as part of any planned activity. Without any doubt, the results of such basic, comparative research on ecosystems will have positive effects on the effective use and conservation of resources.